

COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-276-852-35

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 3
US-08-133-011-107/c
Sequence 107, Application US/08133011
Patent No. 5658727
GENERAL INFORMATION:
APPLICANT: Kang, Angray
APPLICANT: Barbas, Carlos
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
TITLE OF INVENTION: PHAGEMIDS
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5658727th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/133,011
FILING DATE: 08-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/683,602
FILING DATE: 10-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-133-011-107

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 4
US-08-322-730A-107/c
Sequence 107, Application US/08322730A
Patent No. 5759817
GENERAL INFORMATION:
APPLICANT: Barbas, Carlos
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
TITLE OF INVENTION: PHAGEMIDS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10550 No. 5759817th Torrey Pines Road, suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,730A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
APPLICATION NUMBER: US 07/683,602
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bingham, Douglas A
REGISTRATION NUMBER: 32,457
REFERENCE/DOCKET NUMBER: SCRO707P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-784-2937

TELEFAX: 619-784-9399
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-322-730A-107

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCCGCCAC 84
Db 31 ATGAGAGCCACCCGCCAC 14

RESULT 5

US-08-387-874-79/c
Sequence 79, Application US/08387874
Patent No. 5770356

GENERAL INFORMATION:

APPLICANT: Light, Paul L., II

APPLICANT: Lerner, Richard A.

TITLE OF INVENTION: PHAGEMIDS COEXPRESSING A SURFACE

TITLE OF INVENTION: RECEPTOR AND A SURFACE HETEROLOGOUS PROTEIN

NUMBER OF SEQUENCES: 97

CORRESPONDENCE ADDRESS: The Scripps Research Institute, Office of

ADDRESSEE: Patent Counsel

STREET: 10666 No. 5770356th Torrey Pines Road, Suite 220,

STREET: Mail Drop 1PC8

CITY: La Jolla

STATE: CA

COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/387,874

FILING DATE: 22-FEB-1995

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: PCT/US93/08364

APPLICATION NUMBER: 03-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/941,369

FILING DATE: 04-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 303.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937

TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 48 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-387-874-79

Query Match 4.3%; Score 18; DB 1; Length 48;

Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCCGCCAC 84
Db 31 ATGAGAGCCACCCGCCAC 14

RESULT 6

US-08-899-575-35/c

Sequence 35, Application US/08899575
Patent No. 5770440

GENERAL INFORMATION:

APPLICANT: Burton, Dennis R

APPLICANT: Barbas, Carlos F

APPLICANT: Lerner, Richard A

TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 170

CORRESPONDENCE ADDRESS: The Scripps Research Institute, Office of

ADDRESSEE: Patent Counsel

STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,

STREET: Mail Drop 1PC8

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,575

FILING DATE: 24-JUL-1997

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: US 08/276,852

APPLICATION NUMBER: 18-JUL-1994

FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/954,148

FILING DATE: 30-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: SCR1452P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937

TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 48 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-899-575-35

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCCGCCAC 84
Db 31 ATGAGAGCCACCCGCCAC 14

RESULT 7

US-08-899-575-35/c
; Sequence 35, Application US/08899575
; Patent No. 580440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-899-575-35

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 8
US-08-383-619-107/c
; Sequence 107, Application US/08383619
; Patent No. 5955341
; GENERAL INFORMATION:
; APPLICANT: Kang, Angray
; APPLICANT: Barbas, Carlos
; APPLICANT: Lerner, Richard
; TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
; TITLE OF INVENTION: PHAGEMIDS

NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DOUGLAS A. BINGHAM
; STREET: 11300 Sorrento Valley Road, Suite 200
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,619
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,680
; FILING DATE:
; APPLICATION NUMBER: US/07/663,602
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bingham, Douglas A.
; REGISTRATION NUMBER: 32,457
; REFERENCE/DOCKET NUMBER: SCR0371P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1555
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-383-619-107

Query Match 4.3%; Score 18; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 9
US-08-888-366-36/c
; Sequence 36, Application US/0888366
; Patent No. 5972656
; GENERAL INFORMATION:
; APPLICANT: Lopez, Osvaldo
; APPLICANT: Wylie, Dwane E.
; APPLICANT: Wagner, Fred W.
; TITLE OF INVENTION: Mercury Binding Polypeptides and Nucleotides Coding Therefore
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 90 South 7th Street, 3100 No. 5972656west Ctr.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,366
; FILING DATE: 03-JUL-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/187,407
FILING DATE: 27-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,542
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/493,299
FILING DATE: 14-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/324,392
FILING DATE: 14-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.39USC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-888-366-36

Query Match 4.3%; Score 18; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 10
US-08-767-128-43/C
Sequence 43, Application US/08767128
Patent No. 611079
GENERAL INFORMATION:
APPLICANT: WYLLIE, DWANE E.
APPLICANT: LOPEZ, OSVALDO
APPLICANT: MURRAY, PETER JOSEPH
APPLICANT: GOBEL, PETER
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND
NUMBER OF INVENTION: NUCLEOTIDES CODING THEREFORE
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 611079west Center, 30 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 04-DEC-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09258
FILING DATE: 05-JUN-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/541,373
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,798
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.49USF1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/371-5278
TELEFAX: 612/332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-767-128-43

Query Match 4.3%; Score 18; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 11
US-08-907-739-107/C
Sequence 107, Application US/08907739
Patent No. 6235469
GENERAL INFORMATION:
APPLICANT: Kang, Angrey
APPLICANT: Barbos, Carlos
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
NUMBER OF INVENTION: PHAGETIDS
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
STREET: 10666 No. 6235469th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/907,739
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/133,011
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-907-719-107

Query Match 4.3%; Score 18; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 12
US-08-972-564-19/C
Sequence 19, Application US/08972564
Patent No. 6376170
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R.
APPLICANT: Burioni, Roberto
APPLICANT: Williams, R. Anthony
APPLICANT: Sanna, Pietro Paolo
TITLE OF INVENTION: LIGAND CAPTURE-DIRECTED SELECTION OF ANTIBODY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Juba & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,564
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/316,914
FILING DATE: October 3, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Hilde, Ph.D.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD-3809
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS

LOCATION: 1..48
US-08-972-564-19

Query Match 4.3%; Score 18; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 13
US-09-729-597-107/C
Sequence 107, Application US/09729597
Patent No. 6468738
GENERAL INFORMATION:
APPLICANT: Kang, Angray
Barbas, Carlos
Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
PHAGEMIDS
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
Patent Counsel
STREET: 10666 No. 6468738th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/729,597
FILING DATE: 04-Dec-2000
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/133,011
FILING DATE: 1994-09-29
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 107:
US-09-729-597-107

Query Match 4.3%; Score 18; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 14
PCT-US93-08364-79/c
Sequence 79: Application PC/TUS9308364
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PHAGEMIDS COEXPRESSING A SURFACE
RECEPTOR AND A SURFACE HETEROLOGOUS PROTEIN
NUMBER OF SEQUENCES: 97
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08364
FILING DATE: 03-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/941,369
FILING DATE: 04-SEP-1992
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US93-08364-79

Query Match 4.3%; Score 18; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 15
PCT-US95-00067-21/c
Sequence 21: Application PC/TUS9500067
GENERAL INFORMATION:
APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO HERPES
SIMPLEX VIRUS AND METHODS THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00067
FILING DATE: 04-JAN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: PD-3229
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..48
PCT-US95-00067-21

Query Match 4.3%; Score 18; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||
Db 31 ATGAGAGCCACCGCCAC 14

Search completed: March 22, 2004, 05:13:49
Job time: 84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 22, 2004, 03:25:06 / Search time 311 Seconds
(without alignments)
4959.958 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417
Sequence: 1 atgsgggaactgtgcgcag.....gctctctgtaaccacgagn 417

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2438257 seqs, 1849576744 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1556246

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCRT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCRTS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09 NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	4.3	48	13	US-10-079-802-19
C 2	18	4.3	48	15	US-10-016-986-35
C 3	18	4.3	48	15	US-10-273-973-107
C 4	17	4.1	60	10	US-09-908-975-11168
C 5	16	3.8	22	14	US-10-321-195-12
C 6	15	3.6	17	14	US-10-238-700-3063
C 7	15	3.6	20	14	US-10-331-907-358
C 8	15	3.6	25	14	US-10-038-263B-98776
C 9	15	3.6	65	10	US-09-908-975-1082
C 10	15	3.6	65	10	US-09-908-975-28907
C 11	15	3.6	94	9	US-09-864-761-20254
C 12	14	3.4	17	14	US-10-238-700-3062
C 13	14	3.4	17	14	US-10-061-201-809
C 14	14	3.4	17	14	US-10-061-201-810
C 15	14	3.4	17	14	US-10-061-201-811

16	14	3.4	17	14	US-10-061-201-812	Sequence 812, App
C 17	14	3.4	18	14	US-10-181-603-14	Sequence 14, Appl
C 18	14	3.4	20	10	US-09-953-611-85	Sequence 85, Appl
C 19	14	3.4	20	15	US-10-349-143-5592	Sequence 592, Ap
C 20	14	3.4	21	14	US-10-010-920-60	Sequence 60, Appl
C 21	14	3.4	21	14	US-10-008-721-60	Sequence 60, Appl
C 22	14	3.4	24	9	US-09-779-879A-29	Sequence 29, Appl
C 23	14	3.4	24	9	US-09-779-880A-29	Sequence 29, Appl
C 24	14	3.4	24	9	US-09-988-899-51	Sequence 51, Appl
C 25	14	3.4	24	9	US-09-910-120-18	Sequence 18, Appl
C 26	14	3.4	24	10	US-09-833-041-42	Sequence 42, Appl
C 27	14	3.4	24	11	US-09-833-245-42	Sequence 42, Appl
C 28	14	3.4	24	12	US-10-649-273-39	Sequence 39, Appl
C 29	14	3.4	24	12	US-10-351-891-18	Sequence 18, Appl
C 30	14	3.4	24	12	US-10-651-722-39	Sequence 39, Appl
C 31	14	3.4	24	13	US-10-039-785-12	Sequence 12, Appl
C 32	14	3.4	24	13	US-10-153-064-39	Sequence 39, Appl
C 33	14	3.4	24	14	US-10-077-023-127	Sequence 127, App
C 34	14	3.4	24	14	US-10-075-846-44	Sequence 44, Appl
C 35	14	3.4	24	14	US-10-056-884-37	Sequence 37, Appl
C 36	14	3.4	24	14	US-10-080-980-34	Sequence 34, Appl
C 37	14	3.4	24	14	US-10-092-135-44	Sequence 44, Appl
C 38	14	3.4	24	14	US-10-086-185-64	Sequence 64, Appl
C 39	14	3.4	24	14	US-10-081-775-34	Sequence 34, Appl
C 40	14	3.4	24	14	US-10-067-443-39	Sequence 39, Appl
C 41	14	3.4	24	14	US-10-104-943-100	Sequence 100, Appl
C 42	14	3.4	24	14	US-10-120-604-146	Sequence 146, App
C 43	14	3.4	24	14	US-10-067-649-60	Sequence 60, Appl
C 44	14	3.4	24	14	US-10-067-800-29	Sequence 29, Appl
C 45	14	3.4	24	14	US-10-133-797-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-10-079-802-19/C
Sequence 19, Application US/10079802
Publication No. US20020168629A1
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R.
Burton, Roberto
Williamson, R. Anthony
Sanna, Pietro Paolo
TITLE OF INVENTION: LIGAND CAPTURE-DIRECTED SELECTION OF ANTIBODY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Spensley Horn Juba & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/079,802
FILING DATE: 19-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/316,914
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: LISA A. HALL, PH.D.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD-3809
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 19:

```

: SEQUENCE CHARACTERISTICS:
:     LENGTH: 48 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:     MOLECULE TYPE: DNA (genomic)
:     FEATURE:
:         NAME/KEY: CDS
:         LOCATION: 1..48
:     SEQUENCE DESCRIPTION: SEQ ID NO: 19
US-10-079-802-19

```

Query Match	4.3%	Score 18;	DB 13;	Length 48;
Best Local Similarity	100.0%	Pred. No. 37;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||
Db 31 ATGAGAGCCACCGCCAC 14

```

US-10-016-986-35/C
; Sequence 35, Application US/10016986
; Publication No. US20030187247A1
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbos, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 313.2CON1
; CURRENT APPLICATION NUMBER: US/10/016,986
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 09/149,898
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: US 08/999,575
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 08/276,852
; PRIOR FILING DATE: 1994-07-18
; PRIOR APPLICATION NUMBER: US 08/178,302
; PRIOR FILING DATE: 1994-01-06
; PRIOR APPLICATION NUMBER: PCT/US93/09328
; PRIOR FILING DATE: 1993-09-30
; PRIOR APPLICATION NUMBER: US 07/954,148
; PRIOR FILING DATE: 1992-09-30
; NUMBER OF SEQ. ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
; US-10-016-986-35

```

Query Match 4.3%; Score 18; DB 14; Length 48;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY		67	ATGAGAGCCACCGCCAC	84
Db		31	ATGAGAGCCACCGCCAC	14

RESULT 3
US-10-273-973-107/c
; Sequence 107, Application US/10273973
; Publication No. US2004020057A1
; GENERAL INFORMATION:
; APPLICANT: Kang, Andrew
; Barbas, Carlos
; Bernier, Richard A.

TITLE OF INVENTION: HETEROODIMERIC RECEPTOR LIBRARIES USING PHAGEMIDS
 NUMBER OF SEQUENCES: 161
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: The Scripps Research Institute, Office of Patent Counsel
 STREET: 10666 No. US20040002057a1th Torrey Pines Road, TPC-8
 CITY: La Jolla
 STATE: CA
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/273,973
 FILING DATE: 22-Jan-2003
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/133,011
 FILING DATE: 08-JUN-1994
 APPLICATION NUMBER: US 07/683,602
 FILING DATE: 10-APR-1991
 APPLICATION NUMBER: US 07/826,623
 FILING DATE: 27-JAN-1992
 APPLICATION NUMBER: PCT/US 92/03091
 FILING DATE: 10-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Fitting, Thomas
 REGISTRATION NUMBER: 34,163
 REFERENCE/DOCKET NUMBER: SCRF 238.2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-554-2937
 TELEFAX: 619-554-6312
 INFORMATION FOR SEQ ID NO: 107:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 48 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 SEQUENCE DESCRIPTION: SEQ ID NO: 107:
 GS-10-273-973-107

```

Query Match      4.3%; Score 18; DB 15; Length 48;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

QY	67	ATGAGAGCCACCGCCAC	84
Db	31	ATGAGAGCCACCGCCAC	14

```

RESULT 4
US-09-908-975-11168
; Sequence 11168, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Ilat
; APPLICANT: FAIGLER, Simchon
;
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724

```

; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11168
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-11168

Query Match 4.1%; Score 17; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 GAACCTCAGGCTCCTGG 406
DB 37 GAACCTCAGGCTCCTGG 53

RESULT 5

US-10-321-195-12
; Sequence 12, Application US/10321195
; Publication No. US20030148353A1
; GENERAL INFORMATION:
; APPLICANT: Borrebaeck, Carl A.
; APPLICANT: Malmberg-Hager, Ann-Christina
; APPLICANT: Furebring, Christina
; APPLICANT: Soderlund, Ulf H.
; APPLICANT: Ottosson, Rebecca I.
; TITLE OF INVENTION: A Method For In Vitro Molecular
; FILE REFERENCE: 0046-P02150US1
; CURRENT APPLICATION NUMBER: US/10/321,195
; PRIOR APPLICATION NUMBER: 09/445,649
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/GB98/01757
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: GB9712512.4
; PRIOR FILING DATE: 1997-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-321-195-12

Query Match 3.8%; Score 16; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GGAGAGCCACCGCCAC 84
DB 1 GGAGAGCCACCGCCAC 16

RESULT 6

US-10-238-700-3063
; Sequence 3063, Application US/10238700
; Publication No. US20030153521A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: McSwigen, James
; TITLE OF INVENTION: Nucleic Acid Treatment of Diseases or Conditions Related to Level
; FILE REFERENCE: 400/057 (MB1801-1158-A)
; CURRENT APPLICATION NUMBER: US/10/238,700
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: PCT/US 02/16840
; PRIOR FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: US 60/318,471
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 4666
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3063
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-238-700-3063

Query Match 3.6%; Score 15; DB 14; Length 17;
Best Local Similarity 93.3%; Pred. No. 1.6e+03;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 194 ACAGCGCCATGCGGG 208
DB 1 ACAGCGCCATGCGGG 15

RESULT 7

US-10-331-907-358/C
; Sequence 358, Application US/10331907
; Publication No. US20030181660A1
; GENERAL INFORMATION:
; APPLICANT: Todd, John A.
; Hees, John W.
; Caskey, Charles T.
; Cox, Roger D.
; Gerhold, David
; Hammond, Holly
; Hey, Patricia
; Kawaguchi, Yoshihiko
; Merriman, Tony R.
; Metzker, Michael L.
; TITLE OF INVENTION: No. US20030181660A1e1 LDL-Receptor
; NUMBER OF SEQUENCES: 455
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and Vanderhye
; STREET: 1100 No. US20030181660A1e1 Glebe Road, Eighth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: VA 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/331,907
; FILING DATE: 31-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,923A
; FILING DATE: 14-Feb-2001
; APPLICATION NUMBER: PCT/GB98/01102
; FILING DATE: 15-APR-1998
; APPLICATION NUMBER: US 60/043,553
; FILING DATE: 15-APR-1997
; APPLICATION NUMBER: US 60/048,740
; FILING DATE: 05-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: B.V. Sadoff
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4091
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 358:
US-10-331-907-358

Query Match 3.6%; Score 15; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 GGAGGTGCTGCACAG 179
DB 20 GGAGTCTGCACAG 6

RESULT 8
US-10-098-263B-98776

Sequence 98776, Application US/10098263B
Publication No. US20030104410A1
GENERAL INFORMATION:
APPLICANT: Miteam, Michael
TITLE OF INVENTION: Human Microarray
FILE REFERENCE: 3118.1
CURRENT APPLICATION NUMBER: US/10/098,263B
CURRENT FILING DATE: 2003-01-08
PRIOR APPLICATION NUMBER: 60/276,759
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 131066
SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
SEQ ID NO 98776
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapien
US-10-098-263B-98776

Query Match 3.6%; Score 15; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 CTTCTGAGAGCCT 275
DB 9 CTTCTGAGAGCCT 23

RESULT 9
US-09-908-975-1082

Sequence 1082, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1082
LENGTH: 65
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-908-975-1082

Query Match 3.6%; Score 15; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 ACGCTGATCTGCC 107

DB 11 ACGCTGATCTGCC 25

RESULT 10
US-09-908-975-28907

Sequence 28907, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 28907
LENGTH: 65
TYPE: DNA
ORGANISM: Mus musculus
US-09-908-975-28907

Query Match 3.6%; Score 15; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 GAGCTGAAGTTCCA 284
DB 26 GAGCTGAAGTTCCA 40

RESULT 11
US-09-864-761-20254

Sequence 20254, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Autonomax Sequence Listing Engine vers. 1.1
SEQ ID NO 20254
LENGTH: 94
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007097.2
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
OTHER INFORMATION: NT HIT: AB011536.1, EVALUATE 6.60e-01
OTHER INFORMATION: EST_HUMAN HIT: BE677194.1, EVALUATE 5.00e-46
US-09-864-761-20254

Query Match 3.6%; Score 15; DB 9; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 GTGCTGGCCGCGTG 156
DB 49 GTGCTGGCCGCGTG 63

RESULT 12
US-10-238-700-3062
Sequence 3062, Application US/10238700
Publication No. US20030153521A1
GENERAL INFORMATION:
APPLICANT: McSwigen, James
TITLE OF INVENTION: Nucleic Acid Treatment of Diseases or Conditions Related to Level
FILE REFERENCE: 400/057 (MHB01-1158-A)
CURRENT APPLICATION NUMBER: US/10/238,700
PRIOR FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: PCT/US 02/16840
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US 60/318,471
PRIOR FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 4666
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3062
LENGTH: 17
TYPE: RNA
ORGANISM: Homo sapiens
US-10-238-700-3062

Query Match 3.4%; Score 14; DB 14; Length 17;
Best Local Similarity 92.9%; Pred. No. 5.4e+03;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 194 ACAGCCCATGCGG 207
DB 4 ACAGCCCATGCGG 17

RESULT 13
US-10-061-201-809
Sequence 809, Application US/10061201
Publication No. US20030166229A1
GENERAL INFORMATION:
APPLICANT: Shannon, Mark
TITLE OF INVENTION: HUMAN POSH-LIKE PROTEIN 1
FILE REFERENCE: PB0178
CURRENT APPLICATION NUMBER: US/10/061,201
PRIOR FILING DATE: 2002-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/328,205
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 4162
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 809
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapiens
US-10-061-201-809

Query Match 3.4%; Score 14; DB 14; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CTGCCCCAGCCGCC 115
DB 4 CTGCCCCAGCCGCC 17

RESULT 14
US-10-061-201-810
Sequence 810, Application US/10061201
Publication No. US20030166229A1
GENERAL INFORMATION:
APPLICANT: Shannon, Mark
TITLE OF INVENTION: HUMAN POSH-LIKE PROTEIN 1
FILE REFERENCE: PB0178
CURRENT APPLICATION NUMBER: US/10/061,201
PRIOR FILING DATE: 2002-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

Search completed: March 22, 2004, 05:19:21
Job time : 322 secs

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/328,205
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 4162
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 810
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapiens
US-10-061-201-810

Query Match 3.4%; Score 14; DB 14; Length 17;
Best Local Similarity 100.0%; Pred.No. 5.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CTGCCCGAGCGCC 115
DB 3 CTGCCCGAGCGCC 16

RESULT 15
US-10-061-201-811
Sequence 811, Application US/10061201
Publication No. US20030166229A1
GENERAL INFORMATION:
APPLICANT: Shannon, Mark
TITLE OF INVENTION: HUMAN POSH-LIKE PROTEIN 1
FILE REFERENCE: PB0178
CURRENT APPLICATION NUMBER: US/10/061,201
CURRENT FILING DATE: 2002-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/328,205
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 4162
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 811
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapiens
US-10-061-201-811

Query Match 3.4%; Score 14; DB 14; Length 17;
Best Local Similarity 100.0%; Pred.No. 5.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CTGCCCGAGCGCC 115
DB 2 CTGCCCGAGCGCC 15

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 03:19:00 ; Search time 3013 Seconds
(without alignments)
4132.934 Million cell updates/sec

Title: US-10-032-159a-15

Perfect score: 417
Sequence: 1 atgggggagactgtgcgcgag.....gctcctcgtgtaacccagm 417

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 569298

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	16	3.8	65	CD949815 SAN 244 G
2	16	3.8	87	BF180861
3	15	3.6	62	AT739447 w13f01.x
4	15	3.6	70	AA629219 z173f07.s

Result No.	Score	Query Match Length	ID	Description
C 5	15	3.6	76	AI138663 ta80b10.x
6	15	3.6	79	AA731299 nw86g05.s
7	15	3.6	79	AA877223 cb15f03.s
8	15	3.6	81	BH220418 1006094D1
9	15	3.6	85	AI605313 v187h03.x
10	15	3.6	91	BQ758392 EBma05_SQ
11	15	3.6	100	AA510788 v931h04.x
12	15	3.6	100	BF110926 7a30b10.x
13	15	3.6	100	AA502296 1M0541G24
14	15	3.4	37	A2771602 1M0573A17
15	14	3.4	43	AI014286 am46a02.s
16	14	3.4	48	AI358130 gw13f10.x
17	14	3.4	43	A2786204 2M0031N04
18	14	3.4	50	AI05866 AI05866
19	14	3.4	50	AI06969 AI06969
20	14	3.4	50	AI06972 AI06972
21	14	3.4	50	AI06973 AI06973
22	14	3.4	50	AI06974 AI06974
23	14	3.4	50	AI06976 AI06976
24	14	3.4	50	AI06977 AI06977
25	14	3.4	50	AI06978 AI06978
26	14	3.4	56	AZ801785 2M0060E22
27	14	3.4	57	CG733686 1119158B0
28	14	3.4	60	CNS0290X
29	14	3.4	71	AI990431
30	14	3.4	73	AI990431
31	14	3.4	75	BE737572
32	14	3.4	75	CG709037
33	14	3.4	78	AI628233
34	14	3.4	81	A2919233 100601AFO
35	14	3.4	81	CG733644 1119158A0
36	14	3.4	82	CG733644 1119158A0
37	14	3.4	84	CG710246 1119016F0
38	14	3.4	84	N89309 K9029F Huma
39	14	3.4	88	CG721982 1119069H0
40	14	3.4	90	AA172789 m557c01.r
41	14	3.4	90	CG733649 1119158A0
42	14	3.4	92	BE591436 7h57c04.x
43	14	3.4	92	BZ592694 SALX 0284
44	14	3.4	93	CG709055 1119011F1
45	14	3.4	97	CG712696 1119028F0

ALIGNMENTS

RESULT 1
CD949815/c 65 bp mRNA linear EST 15-JUL-2003
LOCUS SAN 244 GenetAg2 Zea mays cDNA, mRNA sequence.
DEFINITION CD949815
ACCESSION CD949815
VERSION CD949815.1 GT32797575
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
Gramineae; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 65)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2003)
Contact: Genopiante
Genopiante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopiante' (http://www.genopiante.com
and http://genopiante-info.inbio.fr).
location/Qualifiers
1. .65
/organism="Zea mays"

FEATURES
source

ORIGIN

Query Match 3.8%; Score 16; DB 14; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.1e+04; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 TCTAGGAACCTCAGGC 400
|||||
Db 22 TCTAGGAACCTCAGGC 7

RESULT 2
LOCUS BF180861
DEFINITION 601804988B1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4036077 5',
mRNA sequence.
ACCESSION BF180861
VERSION BF180861.1 GI:11059003
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 87)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LAM9310 row: 0 column: 22
High quality sequence stop: 87.
Location/Qualifiers
1..87
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4036077"
/issue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

ORIGIN

Query Match 3.8%; Score 16; DB 10; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.2e+04; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 CGGGGCCATCGCCTTC 264
|||||
Db 57 CGGGGCCATCGCCTTC 72

RESULT 3
LOCUS A1739447
DEFINITION w13f01.x1 NCI_CGAP_Co16 Homo sapiens cDNA clone IMAGE:2390137 3',
62 bp mRNA linear EST 20-DEC-1999

ORIGIN

Query Match 3.6%; Score 15; DB 9; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.2e+04; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 TCCTGGAGAGCCTGA 277
|||||
Db 21 TCCTGGAGAGCCTGA 35

RESULT 4
LOCUS AA629219/c
DEFINITION AA629219 70 bp mRNA linear EST 16-OCT-1997
Zu73f07.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:743653
3' similar to TR:G609419 G609419 CHROMOSOME XII COSMID 8039. ?
mRNA sequence.
ACCESSION AA629219
VERSION AA629219.1 GI:2541606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 70)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kritzman, D., Kucada, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Thelings, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project

ORIGIN

Query Match 3.6%; Score 15; DB 9; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.2e+04; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 TCCTGGAGAGCCTGA 277
|||||
Db 21 TCCTGGAGAGCCTGA 35

RESULT 4
LOCUS AA629219/c
DEFINITION AA629219 70 bp mRNA linear EST 16-OCT-1997
Zu73f07.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:743653
3' similar to TR:G609419 G609419 CHROMOSOME XII COSMID 8039. ?
mRNA sequence.
ACCESSION AA629219
VERSION AA629219.1 GI:2541606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 70)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kritzman, D., Kucada, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Thelings, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project

ORIGIN

Query Match 3.6%; Score 15; DB 9; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.2e+04; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 TCCTGGAGAGCCTGA 277
|||||
Db 21 TCCTGGAGAGCCTGA 35

RESULT 4
LOCUS AA629219/c
DEFINITION AA629219 70 bp mRNA linear EST 16-OCT-1997
Zu73f07.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:743653
3' similar to TR:G609419 G609419 CHROMOSOME XII COSMID 8039. ?
mRNA sequence.
ACCESSION AA629219
VERSION AA629219.1 GI:2541606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 70)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kritzman, D., Kucada, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Thelings, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project

JOURNAL

Unpublished (1997)

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40m3 fwd. BT from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

1..70
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="GDB:5930947"
/db_xref="taxon:9606"
/clone="IMAGE:743653"
/sex="male"
/lab_host="DH10B"
/clone_1lb="Soares testis_NHT"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer 15'.
TGTACCATCTGAAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'.
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cots, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 3.6%; Score 15; DB 9; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 262 TTCTGTGAGAGCGCTG 276

Db 60 TTCTGTGAGAGCGCTG 46

^

RESULT 5

A1318663

LOCUS tab0b10.x1 NCI CGAP HSC2 Homo sapiens cDNA clone IMAGE:2050363 3'
DEFINITION similar to TR:Q14489 Q14489 RIBOSOMAL PROTEIN S10; mRNA sequence.

ACCESSION A1318663

VERSION A1318663.1 GI:4034598

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 76)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaapb-r@mail.nih.gov
Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.cDNA Library Preparation: David B. Kitzman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdtp/image/image.html

www-bio.llnl.gov/bdtp/image/image.html

Trace considered overall poor quality

Trace considered overall poor quality

FEATURES

source

Insert Length: 170 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers

1..76
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2050363"
/tissue_type="stem cell 34+/38+"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="NCI CGAP HSC2"
/note="Organ: bone marrow; Vector: PAMP1; mRNA made from bone marrow, stem cells 34+/38+, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 400 bp. Primary library, non-amplified."

ORIGIN

Query Match 3.6%; Score 15; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 GTGCTGTGCCAGCTG 156

Db 36 GTGCTGTGCCAGCTG 50

RESULT 6

AAT731299

LOCUS tw68g05.s1 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:1251800 3'
DEFINITION similar to TR:Q13425 Q13425 BETA2-SYNTROPHIN; mRNA sequence.

ACCESSION AAT731299

VERSION AAT731299.1 GI:2752503

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 79)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaapb-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., David Allman,
Ph.D., Gerald Marcil, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdtp/image/image.html

www-bio.llnl.gov/bdtp/image/image.html

Trace considered overall poor quality

Insert Length: 1095 Std Error: 0.00

Seq primer: -40m3 fwd. BT from Amersham

High quality sequence stop: 1.
Location/Qualifiers1..79
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1251800"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_1lb="NCI CGAP GCB1"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was prepared from human consillar cells enriched for germinal center B cells by flow sorting (CD20+, IGD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CSER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTACCACTCTAGAGGAGGCGCCCTCATTTTCTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 3.6%; Score 15; DB 9; Length 79;

Best Local Similarity 100.0%; Pred. No. 3.3e+04;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 125 ACCTGCCGAGGCCA 139

Db 36 ACCTGCCGAGGCCA 50

RESULT 7

AA877223/c

LOCUS

DEFINITION ob15f03.s1 NCI CGAP Kid3 Homo sapiens cDNA IMAGE:1323773 3'

Similar to TR:000554 000554 P21-ARC. [1] ; mRNA sequence.

AA877223

VERSION AA877223.1 GI:2986300

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

79 bp mRNA linear EST 25-MAR-1998
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: www-bio.1nl.gov/bdrip/image/image.html

Trace considered overall poor quality
Seq primer: -40m3 fwd. RT from Amersham
High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
1..79
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1323773"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pRTT3-Pac (Pharmacia) with a modified polylinker, Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match

3.6%; Score 15; DB 9; Length 79;

Best Local Similarity 100.0%; Pred. No. 3.3e+04;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 262 TTCCTGAGAGCCTG 276

Db 63 TTCCTGAGAGCCTG 49

RESULT 8

BH220418

LOCUS

DEFINITION 1006094D10.x1 1006 - RescuemU Grid G Zea mays genomic, genomic survey sequence.

ACCESSION BH220418

VERSION BH220418.1 GI:16814677

KEYWORDS

SOURCE

ORGANISM

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 81)

Walbot, V.

Unpublished (2001)

JOURNAL

COMMENT

Maize genomic sequences found using engineered RescuemU transposon
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site so sequence was trimmed. Post-ligation sequence submitted separately.
Plate: 1006094 row: 28
Class: transposon-tagged.

FEATURES

source

Location/Qualifiers
1..81
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="1006 - RescuemU Grid G"
/note="Organ: leaf; Vector: RescuemU (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII; RescuemU is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescuemU, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescuemU'. Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 3.6%; Score 15; DB 28; Length 81;

Best Local Similarity 100.0%; Pred. No. 3.4e+04;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 GATGATGAGAGGCCA 77

Db 25 GATGATGAGAGGCCA 39

RESULT 9

AL605313/c

LOCUS

DEFINITION vmerh03.x1 StrataGene mouse heart (4937316) Mus musculus cDNA clone

IMAGE:1038965 3', mRNA sequence.
 ACCESSION A1605313
 VERSION A1605313.1 GI:4614480
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 85)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:578989
 This clone was previously sequenced on the 5' end only, this new data is from the 3' end.
 Location/Qualifiers
 1..85
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NIH SWISS"
 /db_xref="taxon:10090"
 /clone="IMAGE:1038965"
 /sex="pooled"
 /tissue_type="heart"
 /dev_stage="13 day embryos"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_1lb="Stratagene mouse heart (#937316)"
 /note="Organ: heart; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'".
 ORIGIN
 Query Match 3.6%; Score 15; DB 9; Length 85;
 Best Local Similarity 100.0%; Pred. No. 3.4e+04;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 211 GGGCATTGCTGGAT 225
 |||||
 |||||
 Db 37 GGGCATTGCTGGAT 23
 RESULT 10
 BQ758392 91 bp mRNA linear EST 26-JUL-2002
 EBMa05_S0001_108 R maternal, 12 DPA, no treatment, cv Optic, EBMa05
 Hordeum vulgare subsp. vulgare cDNA clone EBMa05_S0001_108 5', mRNA
 sequence.
 ACCESSION BQ758392
 VERSION BQ758392.1 GI:21966864
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poidea; Triticeae; Hordeum.
 1 (bases 1 to 91)
 Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L., Ramsay,L., Machray,G., Marshall,D.F.M. and Vaughn,R.

TITLE Development of Barley Transcriptome Resources
 JOURNAL Unpublished (2001)
 COMMENT Contact: Waugh R, Marshall DF
 Genome Dynamics/Computational Biology
 Scottish Crop Research Institute
 Invergowrie, Dundee, DD2 5DA, Scotland, UK
 Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: estecr1.scri.ac.uk.
 Location/Qualifiers
 1..91
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Optic"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="EBMa05_S0001_108"
 /tissue_type="maternal"
 /dev_stage="12 DPA"
 /lab_host="DH10B"
 /clone_1lb="maternal, 12 DPA, no treatment, cv Optic, EBMa05"
 /note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I; Non-normalised library, directionally cloned into pSPORT1. Derived from maternal tissue dissected from developing grains (12 days post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BSRG/SBRAD funded cereal IGF (Investigating Gene Function) project."
 ORIGIN
 Query Match 3.6%; Score 15; DB 13; Length 91;
 Best Local Similarity 100.0%; Pred. No. 3.5e+04;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 160 GAGGAGGAGCTGCTG 174
 |||||
 |||||
 Db 38 GAGGAGGAGCTGCTG 52
 RESULT 11
 AA510788 100 bp mRNA linear EST 08-JUL-1997
 AA510788
 LOCUS Vg31h04.r1 Soares mammary_gland_MOMG Mus musculus cDNA clone
 DEFINITION IMAGE:863191 5', mRNA sequence.
 ACCESSION AA510788
 VERSION AA510788
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 100)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:507279
 Seg primer: -28n3 rev2 ET from Amersham
 High quality sequence stop: 84.
 Location/Qualifiers

FEATURES

ORIGIN

10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

Query Match 3.6%; Score 15; DB 28; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.5e+04;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 GGACGAGGAGAGGT 170
Db 73 GGACGAGGAGAGGT 87

RESULT 14
AZ771602 37 bp DNA linear GSS 16-FEB-2001
LOCUS 1M0573A17R Mouse 10kb plasmid UGCG1M library Mus musculus genomic
DEFINITION clone UGCG1M0573A17 R, genomic survey sequence.
ACCESSION AZ771602
VERSION AZ771602.1 GI:12893864
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 37)

TITLE
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islem, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, W., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
COMMENT Unpublished (2000)

Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0573 row: A column: 17
Seq primer: CACACAGGAAACGCTATGACC
Clas: plasmid ends

High quality sequence stop: 37.
Location/Qualifiers

FEATURES

source

1..37
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG1M0573A17"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCG1M library"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/shares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to

ORIGIN

10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

Query Match 3.4%; Score 14; DB 28; Length 37;
Best Local Similarity 100.0%; Pred. No. 8.3e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 AGGTGCTGCACAGC 180
Db 9 AGGTGCTGCACAGC 22

RESULT 15
A1014286 43 bp mRNA linear EST 15-JUN-1998
LOCUS 1M46602.1 Johnston frontal cortex Homo sapiens cDNA clone
DEFINITION IMAGE:1538570 3' similar to gb:U87789 IG GAMMA-1 CHAIN C REGION (HUMAN); mRNA sequence.
ACCESSION A1014286
VERSION A1014286.1 GI:3228118
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 43)

TITLE
Haller, J., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, K., Martin, J., Moore, B., Scheinberg, R., Steptoe, M., Tan, F., Treising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
Washu-NCI human EST Project

JOURNAL
COMMENT Unpublished (1997)

Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

This clone is available royalty-free through LIND; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -40m13 fwd. RT from Amersham

High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

1..43
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1538570"
/sex="Male"
/tissue_type="pooled frontal lobe"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Johnston frontal cortex"
/note="Organ: Brain; Vector: Bluescript SK-; Site: 1; EcORI; Stanley Neuropathology Consortium (www.stanleylab.org) brains S-58, S-65, S-67, S-78. Random + oligo-dT primed into EcORI site of Zap II Vector. Mass excised. Avg insert length 1.9kb. Custom library provided by Dr. Nancy Johnston [(410) 614-3918, nlj@wustl.welch.jnu.edu].

ORIGIN

Query Match 3.4%; Score 14; DB 9; Length 43;
Best Local Similarity 100.0%; Pred. No. 8.5e+04;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 300 CACCTGTGTCACCG 313
| | | | | | | | | |
Db 21 CACCTGTGTCACCG 34

Search completed: March 22, 2004, 05:12:28
Job time : 3031 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 03:12:56 ; Search time 3050 Seconds

(without alignments)
5925.916 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417
Sequence: 1 atgggggagactgctgcgcag.....gctcctcgtgtaacccagm 417

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 2167151695 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1846492

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ses:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_hum:*
17: em_hun:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ses:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rdt:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	18	4.3	39	6	A46185	A46185 Sequence 23
C 2	18	4.3	39	6	AR351110	AR351110 Sequence
C 3	18	4.3	48	6	AR038280	AR038280 Sequence
C 4	18	4.3	48	6	AR075276	AR075276 Sequence
C 5	18	4.3	48	6	AR081918	AR081918 Sequence
C 6	18	4.3	48	6	AR108691	AR108691 Sequence
C 7	18	4.3	48	6	AR152688	AR152688 Sequence
C 8	18	4.3	48	6	IS8569	IS8569 Sequence 35
C 9	18	4.3	48	6	IS1300	IS1300 Sequence 10
C 10	18	4.3	48	6	AR238528	AR238528 Sequence
C 11	18	4.3	48	6	AR357623	AR357623 Sequence
C 12	18	4.3	48	6	AX323341	AX323341 Sequence
C 13	17	4.1	66	6	E29068	E29068 Modified in
C 14	16	3.8	22	6	A81881	A81881 Sequence 12
C 15	16	3.8	22	6	AR120684	AR120684 Sequence
C 16	16	3.8	22	6	AR266660	AR266660 Sequence
C 17	16	3.8	22	6	BD135550	BD135550 A method
C 18	16	3.8	81	9	HSXTCRA07	X83782 H.sapiens m
C 19	15	3.6	20	6	A91597	A91597 Sequence 12
C 20	15	3.6	20	6	AR163855	AR163855 Sequence
C 21	15	3.6	20	6	AR305404	AR305404 Sequence
C 22	15	3.6	20	6	AR309508	AR309508 Sequence
C 23	15	3.6	20	6	BD023379	BD023379 Method fo
C 24	15	3.6	20	6	BD106315	BD106315 Novel LDL
C 25	15	3.6	29	6	BD187047	BD187047 Polypepti
C 26	15	3.6	42	6	AR336828	AR336828 Sequence
C 27	15	3.6	42	6	AR336829	AR336829 Sequence
C 28	15	3.6	42	9	HSAR10886	AJ010866 Homo sapi
C 29	15	3.6	71	6	AR316846	AR316846 Sequence
C 30	15	3.6	71	6	AR336826	AR336826 Sequence
C 31	15	3.6	71	6	AX039943	AX039943 Sequence
C 32	15	3.6	74	11	AB091414	AB091414 Sus scrofa
C 33	15	3.6	82	10	RNO230619	AJ230619 Rattus no
C 34	15	3.6	99	6	AX023606	AX023606 Sequence
C 35	14	3.4	17	6	AX531300	AX531300 Sequence
C 36	14	3.4	17	6	AX531301	AX531301 Sequence
C 37	14	3.4	17	6	AX531302	AX531302 Sequence
C 38	14	3.4	17	6	AX531303	AX531303 Sequence
C 39	14	3.4	18	6	AR054201	AR054201 Sequence
C 40	14	3.4	18	6	AR087097	AR087097 Sequence
C 41	14	3.4	18	6	AR121118	AR121118 Sequence
C 42	14	3.4	18	6	AX754814	AX754814 Sequence
C 43	14	3.4	18	6	AX754824	AX754824 Sequence
C 44	14	3.4	18	6	AX754837	AX754837 Sequence
C 45	14	3.4	20	6	AR294257	AR294257 Sequence

ALIGNMENTS

RESULT 1
A46185/c 39 bp DNA linear PAT 07-MAR-1997
LOCUS
DEFINITION Sequence 23 from Patent WO951914.
ACCESSION A46185
VERSION A46185.1 GI:2300433
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.

REFERENCE
AUTHORS Sodoyer, R., Aujaume, L. and Geoffroy, F.
TITLE METHOD FOR PREPARING A MULTICOMBINATORIAL LIBRARY OF ANTIBODY GENE
EXPRESSION VECTORS
JOURNAL Patent: WO 9521914-A 23 17-AUG-1995;

Pred. No. is the number of results predicted by chance to have a

COMMENT PASTEUR MERIEUX SERUMS VACC (FR)
Other publication AU 166895 950829
Other publication FR 2715940 950811.
FEATURES
source 1. .39
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 4.3%; Score 18; DB 6; Length 39;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 ATGGAGAGCCACCGCCAC 84
|||||
27 ATGGAGAGCCACCGCCAC 10

RESULT 2
AR361110/c AR361110 39 bp DNA linear PAT 17-AUG-2003
LOCUS AR361110
DEFINITION Sequence 23 from patent US 6595697.
ACCESSION AR361110
VERSION AR361110.1 GI:33768813
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 39)
AUTHORS Sodoyer, R., Aujame, L. and Geoffrey, F.
TITLE Process for preparing a multicombinatorial library of vectors for
expressing antibody genes
JOURNAL Patent: US 6595697-A 23-29-JUL-2003;
FEATURES Location/Qualifiers
source 1. .39
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 4.3%; Score 18; DB 6; Length 39;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 ATGGAGAGCCACCGCCAC 84
|||||
27 ATGGAGAGCCACCGCCAC 10

RESULT 3
AR038280/c AR038280 48 bp DNA linear PAT 29-SEP-1999
LOCUS AR038280
DEFINITION Sequence 35 from patent US 5804440.
ACCESSION AR038280
VERSION AR038280.1 GI:5956997
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 48)
AUTHORS Burton, D.R., Barbas, C.F. and Lerner, R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency
virus
JOURNAL Patent: US 5804440-A 35-08-SEP-1999;
FEATURES Location/Qualifiers
source 1. .48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 ATGGAGAGCCACCGCCAC 84
|||||
Db 31 ATGGAGAGCCACCGCCAC 14

RESULT 4
AR075276/c AR075276 48 bp DNA linear PAT 28-AUG-2000
LOCUS AR075276
DEFINITION Sequence 107 from patent US 5955341.
ACCESSION AR075276
VERSION AR075276.1 GI:10002027
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 48)
AUTHORS Kang, A., Barbas, C. and Lerner, R.
TITLE Heterodimeric receptor libraries using phagemids
JOURNAL Patent: US 5955341-A 107-21-SEP-1999;
FEATURES Location/Qualifiers
source 1. .48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 ATGGAGAGCCACCGCCAC 84
|||||
Db 31 ATGGAGAGCCACCGCCAC 14

RESULT 5
AR081918/c AR081918 48 bp DNA linear PAT 31-AUG-2000
LOCUS AR081918
DEFINITION Sequence 36 from patent US 5972656.
ACCESSION AR081918
VERSION AR081918.1 GI:10008644
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 48)
AUTHORS Lopez, O., Wylie, D.E. and Wagner, F.W.
TITLE Mercury binding polypeptides and nucleotides coding therefore
JOURNAL Patent: US 5972656-A 36-26-OCT-1999;
FEATURES Location/Qualifiers
source 1. .48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 ATGGAGAGCCACCGCCAC 84
|||||
Db 31 ATGGAGAGCCACCGCCAC 14

RESULT 6
AR108691/c AR108691 48 bp DNA linear PAT 14-FEB-2001
LOCUS AR108691
DEFINITION Sequence 43 from patent US 611079.
ACCESSION AR108691
VERSION AR108691.1 GI:12824178
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 48)
AUTHORS Wylie,D.E., Lopez,O., Murray,P,Joseph, and Goebel,P.
TITLE Lead binding polypeptides and nucleotides coding therefore
JOURNAL Patent: US 611079-A 43 29-AUG-2000;
FEATURES Location/Qualifiers

SOURCE

1..48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 7

AR152688/c AR152688 48 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 107 from patent US 6235469.
DEFINITION AR152688
ACCESSION AR152688
VERSION AR152688.1 GI:15120220
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1 (bases 1 to 48)
AUTHORS Kang,A., Barbas,C. and Lerner,R.A.
TITLE Heterodimeric receptor libraries using phagemids
JOURNAL Patent: US 6235469-A 107 22-MAY-2001;
FEATURES Location/Qualifiers

SOURCE

1..48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 8

158569/c 158569 48 bp DNA linear PAT 07-OCT-1997
LOCUS Sequence 35 from patent US 5652138.
DEFINITION 158569
ACCESSION 158569
VERSION 158569.1 GI:2477807
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1 (bases 1 to 48)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5652138-A 35 29-JUL-1997;
FEATURES Location/Qualifiers

SOURCE

1..48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 9

161300/c 161300 48 bp DNA linear PAT 07-OCT-1997
LOCUS Sequence 107 from patent US 5658727.
DEFINITION 161300
ACCESSION 161300
VERSION 161300.1 GI:2479248
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1 (bases 1 to 48)
AUTHORS Barbas,C., Kang,A. and Lerner,R.A.
TITLE Heterodimeric receptor libraries using phagemids
JOURNAL Patent: US 5658727-A 107 19-AUG-1997;
FEATURES Location/Qualifiers

SOURCE

1..48
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 10

AR238528/c AR238528 48 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 107 from patent US 6468738.
DEFINITION AR238528
ACCESSION AR238528
VERSION AR238528.1 GI:27283526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1 (bases 1 to 48)
AUTHORS Kang,A., Barbas,C. and Lerner,R.A.
TITLE Heterodimeric receptor libraries using phagemids
JOURNAL Patent: US 6468738-A 107 22-OCT-2002;
FEATURES Location/Qualifiers

SOURCE

1..48
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 11

AR367623/c AR367623 48 bp DNA linear PAT 12-SEP-2003
LOCUS Sequence 19 from patent US 6376170.
DEFINITION AR367623
ACCESSION AR367623
VERSION AR367623.1 GI:34600987
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.

1 (bases 1 to 48)

REFERENCE Burton,D.R., Burioni,R., Williamson,R.A. and Sanna,P.P.

TITLE Ligand capture-directed selection of antibody

JOURNAL Patent: US 6376170-A 19 23-APR-2002;

FEATURES Location/Qualifiers

source

1. 48

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;

Best Local Similarity 100.0%; Pred.No. 8.8e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84

Db 31 ATGAGAGCCACCGCCAC 14

RESULT 12

AX323341/c

LOCUS AX323341 48 bp DNA linear PAT 07-JAN-2002

DEFINITION Sequence 79 from Patent EP162270.

ACCESSION AX323341

VERSION AX323341.1 GI:18094102

KEYWORDS

SOURCE

unidentified

unidentified

unclassified.

REFERENCE 1

.AUTHORS Light,J.P. and Lerner,R.A.

TITLE Phagemids coexpressing a surface receptor and a surface

JOURNAL Patent: EP 1162270-A 79 12-DEC-2001;

FEATURES THE SCRIPPS RESEARCH INSTITUTE (US)

source

1. 48

/organism="unidentified"

/mol_type="unassigned DNA"

/db_xref="taxon:32644"

/note="synthesized"

ORIGIN

Query Match 4.3%; Score 18; DB 6; Length 48;

Best Local Similarity 100.0%; Pred.No. 8.8e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84

Db 31 ATGAGAGCCACCGCCAC 14

RESULT 13

E29068

LOCUS E29068 66 bp DNA linear PAT 18-JUN-2001

DEFINITION Modified interferon tau-3.

ACCESSION E29068

VERSION E29068.1 GI:13025516

KEYWORDS JP 1999042089-A/3.

SOURCE unidentified

ORGANISM unidentified

unclassified.

REFERENCE 1 (bases 1 to 66)

.AUTHORS Maatko,I. and Takashi,N.

TITLE Modified interferon tau-3

JOURNAL Patent: JP 1999042089-A 3 16-FEB-1999;

COMMENT

OS Unidentified

PN JP 1999042089-A/3

PD 16-FEB-1999

PR 29-JUL-1997 JP 1997203137

PI MASANO ISHIMURA,TAKASHI NISHIGAKI

PC C12N15/09,A61K38/21,C07K7/06,C07K14/555,C12N1/21,C12P21/02//

PC (C12N15/09,C12R1:91),(C12P21/02,C12R1:19),C12N15/00,A61K37/66,

PC (C12N15/00,C12R1:91)

CC Strandedness: Single;

CC Topology: Linear;

FH Key Location/Qualifiers

FT source 1. 66

/organism="Unidentified".

Location/Qualifiers

1. 66

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

ORIGIN

Query Match 4.1%; Score 17; DB 6; Length 66;

Best Local Similarity 100.0%; Pred.No. 2.8e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 GAACCTCAGGCTCTCTGG 406

Db 42 GAACCTCAGGCTCTCTGG 58

RESULT 14

A81881

LOCUS A81881 22 bp DNA linear PAT 21-JAN-2000

DEFINITION Sequence 12 from Patent WO9858080.

ACCESSION A81881

VERSION A81881.1 GI:6731895

KEYWORDS

SOURCE

unidentified

unidentified

unclassified.

REFERENCE 1 (bases 1 to 22)

.AUTHORS Soederlund,U.H. and Borrebaeck,C.A.

TITLE A METHOD FOR IN VITRO MOLECULAR EVOLUTION OF PROTEIN FUNCTION

JOURNAL Patent: WO 9858080-A 12 23-DEC-1998;

FEATURES CRIPPS JOHANNA ELIZABETH (GB); SOEDERLUND Ulf HANS ESKIL (SE)

source

1. 22

/organism="unidentified"

/mol_type="unassigned DNA"

/db_xref="taxon:32644"

ORIGIN

Query Match 3.8%; Score 16; DB 6; Length 22;

Best Local Similarity 100.0%; Pred.No. 1e+04;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GGAGAGCCACCGCCAC 84

Db 1 GGAGAGCCACCGCCAC 16

RESULT 15

AR120684

LOCUS AR120684 22 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 12 from patent US 6159690.

ACCESSION AR120684

VERSION AR120684.1 GI:14104260

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

unclassified.

REFERENCE 1 (bases 1 to 22)

.AUTHORS Borrebaeck,C.Arne,Krister., Soederlund,U.Hans,Eskil. and

TITLE Method for in vitro molecular evolution of protein function using

JOURNAL exonuclease and amplification

PATENT: US 6159690-A 12 12-DEC-2000;

FEATURES Location/Qualifiers

source

1. .22
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 3.8%; Score 16; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GGAGAGCCACCGCCAC 84
|||||
Db 1 GGAGAGCCACCGCCAC 16

Search completed: March 22, 2004, 04:21:44
Job time : 3056 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 02:19:50 ; Search time 391 Seconds
(without alignments)
4530.687 Million cell updates/sec

Title: US-10-032-159A-15
Perfect score: 417
Sequence: 1 atggggggaactctgcgcag.....ggtctctgcaaccacagm 417

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 337863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3485926

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Listing first 45 summaries

Database : N: Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	4.3	33	2	AAQ44650 PCR prime
C 2	18	4.3	39	2	AAQ95082 Multicomb
C 3	18	4.3	48	2	AAQ78813 Membrane
C 4	18	4.3	48	2	AAQ64423 PCR prime
C 5	18	4.3	48	2	AAQ94730 cpiII mem
C 6	18	4.3	48	2	AAQ40878 pcomb2-3 F
C 7	18	4.3	48	2	AAQ58272 M13mp18 P
C 8	18	4.3	48	2	AAQ32124 Primer G-
C 9	18	4.3	48	3	AAQ30998 Primer G-
C 10	18	4.3	48	6	AAQ38725 cpiII mem
C 11	18	4.3	48	8	ADB84272 Dielsidero
C 12	17	4.1	60	6	ABY38420 Human spl
C 13	17	4.1	66	2	AAQ18458 PCR prime
C 14	15	3.6	17	7	ABZ61630 Human H-R
C 15	15	3.6	20	2	AAV85968 Mouse lrp
C 16	15	3.6	20	6	AAV41022 Primer CB
C 17	15	3.6	20	6	ABY94287 Human C/E
C 18	15	3.6	25	8	ACT98785 Human mic
C 19	15	3.6	29	7	ABZ71170 Human HIF
C 20	15	3.6	42	3	AAQ68563 TMOP svnt
C 21	15	3.6	42	9	ADC35520 Trypsin m
C 22	15	3.6	42	9	ADC35521 Trypsin m
C 23	15	3.6	43	3	AAQ65007 Human CTL

24	15	3.6	44	3	AAQ65006 Human CTL
C 25	15	3.6	52	3	AAQ65011 Human CTL
26	15	3.6	53	3	AAQ65010 Human CTL
27	15	3.6	65	6	ABN28334 Rat splic
28	15	3.6	65	6	ABN56159 Mouse spl
29	15	3.6	71	3	AAQ68566 Green flu
30	15	3.6	71	9	ADC35465 Pesticiida
C 31	15	3.6	77	2	AAQ45482 Sense str
32	15	3.6	94	4	AAQ122793 Probe #12
33	15	3.6	94	4	ABQ67880 Human foe
34	15	3.6	94	4	AAQ148095 Probe #16
35	15	3.6	94	4	ABQ49961 Human bre
36	15	3.6	94	4	ABQ49934 Probe #13
37	15	3.6	94	4	AAQ42028 Human bon
38	15	3.6	94	4	AAQ16284 Human bra
39	15	3.6	94	4	ABQ41637 Human l1v
40	15	3.6	94	5	AAQ108467 Probe #84
41	15	3.6	94	6	ABQ16062 Human gen
42	15	3.6	99	3	AAQ36687 RSV and P
C 43	14	3.4	17	6	ABQ90099 Human POS
44	14	3.4	17	6	ABQ90097 Human POS
45	14	3.4	17	6	ABQ90098 Human POS

ALIGNMENTS

RESULT 1
AAQ44650/c
ID AAQ44650 standard; DNM; 33 BP.
XX
AC AAQ44650;
XX
DT 25-MAR-2003 (revised)
DT 04-OCT-1994 (first entry)
XX
DE PCR primer for amplifying mature cpiII membrane anchor gene.
XX
KW Expression; insertion sequence; expression vector; hybrid;
KW filamentous phage; Rd, bacteriophage; cpiIII; membrane anchor; ss.
XX
XX Synthetic.
XX OS
XX EN WC9405781-A1.
XX
PD 17-MAR-1994.
XX
PF 03-SEP-1993; 93MO-US008364.
XX
PR 04-SEP-1992; 92US-00941369.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Light JP, Lerner RA;
XX
XX WPI; 1994-101186/12.
XX
XX Filamentous phage comprising a heterologous polypeptide and a
PT hetero:dimer - is used to detect the presence of a preselected ligand in
a sample.
XX
PS Example 2g; Page 123; 232pp; English.
XX
CC Two primers (AAQ44650, AAQ44651) were used to amplify the cpiII membrane
CC anchor domain using the replicative form of M13mp18 as a template
CC molecule for PCR. The primers also incorporated two restriction enzyme
CC recognition sites (SpeI and NheI) into the amplified molecule. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 33 BP; 5 A; 8 C; 12 G; 8 T; 0 U; 0 Other;
Query Match 4.3%; Score 18; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
 |||||
 Db 31 ATGAGAGCCACCGCCAC 14

RESULT 2
 AAQ95082/c
 ID AAQ95082 standard; DNA; 39 BP.
 XX
 AC AAQ95082;
 XX
 DT 19-MAR-1996 (first entry)
 XX
 DE Multicombinatorial vector library primer pVH GIII*.
 XX
 KM Multicombinatorial library; expression vector; antibody; recombination;
 KM light chain; heavy chain; variable region; pACYC177; lac promoter; attP;
 KM kanamycin resistance; origin of replication; pBluescript; pBR328; attB;
 KM pE1B signal sequence; chloramphenicol resistance; ampicillin resistance;
 KM ColEI; M13 phage; gene III; fusion protein; lambda; recombinase; int;
 KM phagemid; ss.
 XX
 OS Synthetic.
 XX
 PM FR2715940-A1.
 XX
 PD 11-AUG-1995.
 XX
 PF 10-FEB-1994; 94FR-00001519.
 XX
 PR 10-FEB-1994; 94FR-00001519.
 XX
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS.
 XX
 PI Sodoyer R, Aujame L, Geoffroy F;
 XX
 DR WPI; 1995-277038/37.
 XX
 PT Prepn. of multi-combinatorial vector library expressing antibody
 PT fragment(s) - by recombination of vector(s) contg., separately, genes for
 PT light and heavy chain variable region(s).
 XX
 PS Example 2; Page 25; 39pp; French.

The preparation of a multicombinatorial vector library expressing
 fragment of antibodies is carried out by recombining two vectors each
 contg. separately the gene encoding an antibody light or heavy chain
 variable region. The vectors contg the light chain variable region are
 derived from the plasmid pACYC177 and contain the kanamycin resistance
 gene and the p15a origin of replication, the lac promoter from
 pBluescript, the pE1B signal sequence and the chloramphenicol resistance
 gene from pBR328. The gene encoding the light chain variable region is
 inserted in frame with the pE1B signal sequence to produce the plasmid
 pM87. The vector expressing the heavy chain variable region is derived
 from the plasmid pBluescript SKI+ and contains the ampicillin resistance
 gene and the ColEI origin of replication from that plasmid. The plasmid
 has inserted the pE1B signal sequence and the sequence encoding the gene
 III product from M13 phage. The gene encoding the heavy chain variable
 region is inserted in frame with the M13 gene III sequence so as to
 produce a fusion product. The heavy chain vector is designated pM82.
 Both plasmids contain recombination sites: the light chain vector
 contains the lambda attB site whilst the heavy chain vector contains the
 attP site. When introduced into a suitable host cell and contacted by the
 recombinase protein int, the two vector recombine to produce the 8.6 kb
 recombinant phagemid pM85. This vector forms the basis of the library.
 The primers AAQ95082-3 are used to amplify a 663 bp fragment of the
 sequence encoding the gene III protein from the phage M13 mp18 (bases
 2223-2885). The resultant fragment is inserted into pM82 to generate the
 plasmid pM830. The method can be used to produce a library which contains
 a higher number of antibody light and heavy chain clones

SQ Sequence 39 BP; 5 A; 7 C; 14 G; 13 T; 0 U; 0 Other;

Query Match 4.3%; Score 18; DB 2; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
 |||||
 Db 27 ATGAGAGCCACCGCCAC 10

RESULT 3
 AAQ78813/c
 ID AAQ78813 standard; DNA; 48 BP.
 XX
 AC AAQ78813;
 XX
 DT 25-MAR-2003 (revised)
 XX
 DE 19-UTL-1995 (first entry)
 XX
 KM Membrane anchor domain cpIII PCR primer G-3(F).
 XX
 KM Bacteriophage coat protein; membrane anchor domain cpIII; HMCV;
 KM human cytomegalovirus; antibody; immunoassay; immunotherapy; PCR primer;
 KM ss.
 XX
 OS Synthetic.
 XX
 PM WO9425490-A1.
 XX
 PD 10-NOV-1994.
 XX
 PF 29-APR-1994; 94WO-US004705.
 XX
 PR 30-APR-1993; 93US-00055985.
 XX
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Burton DR, Barbas C, Burton R, Williamson A;
 XX
 DR WPI; 1994-358194/44.
 XX
 PT Human monoclonal antibodies (Mabs) against human cytomegalovirus - also
 PT nucleic acids and cell lines producing the Mabs, useful in diagnosis and
 PT immunotherapy.
 XX
 PS Example 1b1; Page 76; 171pp; English.

AAQ78813 and AAQ78814 are a pair of primers for the PCR amplification of
 CC AAQ78812, which encodes AAR62927 bacteriophage coat protein membrane
 CC anchor domain cpIII. It was used in the construction of an expression
 CC vector for a human cytomegalovirus (HCMV) monoclonal antibody (Mab). The
 CC Mab could be used to detect HCMV, and anti-human HCMV Abs in human
 CC patients via a competitive immunoassay. The Mabs may also be useful in
 CC immunotherapy. (Updated on 25-MAR-2003 to correct FN field.)
 XX
 SQ Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;

Query Match 4.3%; Score 18; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
 |||||
 Db 31 ATGAGAGCCACCGCCAC 14

RESULT 4
 AAQ64223/c
 ID AAQ64223 standard; DNA; 48 BP.
 XX
 AC AAQ64223;
 XX

```
DT 25-MAR-2003 (revised)
DT 10-NOV-1994 (first entry)
XX
DE PCR primer G-3(F) to clone membrane anchor domain of M13 cpIII.
XX
XX Human immunodeficiency virus; HIV1; glycoprotein; epitope;
XX neutralisation; monoclonal antibody; light chain; variable region;
XX filamentous phage M13; M13mp18; coat protein cpIII; PCR;
XX polymerase chain reaction amplification; ss.
XX
OS Synthetic.
XX
XX WO9407922-A1.
XX
XX 14-APR-1994.
XX
XX 30-SEP-1993; 93WO-US009328.
XX
XX 30-SEP-1992; 92US-00954148.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Burton DR, Barbas CF, Lerner RA;
XX
XX WPI; 1994-135516/16.
XX
XX New human monoclonal antibodies neutralising HIV - react with gp120 or
XX gp1 and nucleic acid encoding them, useful for in vivo or in vitro
XX diagnosis and for passive immuno-therapy.
XX
XX Example 1; Page 146; 248pp; English.
XX
XX M13mp18 replicative form DNA was used as a template for isolating the
XX gene encoding the membrane anchor domain at cpIII. Two PCR amplifications
XX were performed for construction of a DNA fragment consisting of the
XX mature gene for cpIII membrane anchor domain located 5' to a sequence
XX encoding the lacZ promoter, operator and cap-binding site for controlling
XX light chain expression. Primers G-3(F) and G-3(B) (AA064223 and AA064224,
XX respectively) were used for the first PCR amplification. (Updated on 25-
XX MAR-2003 to correct PN field.)
XX
SQ Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;
Query Match 4.3%; Score 18; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14
RESULT 5
AA094730/c
ID AA094730 standard; DNA; 48 BP.
XX
XX AA094730;
XX
XX 16-OCT-2003 (revised)
XX 22-FEB-1996 (first entry)
XX
XX cpIII membrane anchor gene PCR primer, G-3 (F).
XX
XX Herpes simplex virus; type I; type II; monoclonal antibody; diagnosis;
XX neutralisation; immuno-therapy; ss.
XX
XX Enterobacteria phage M13.
XX
XX WO9518634-A1.
XX
XX 13-JUL-1995.
XX
XX 04-JAN-1995; 95WO-US000067.
XX
XX
```

```
XX
PR 04-JAN-1994; 94US-00178201.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Burton DR, Williamson RA, Burton R, Sarna PP;
XX
XX WPI; 1995-254909/33.
XX
XX Human monoclonal antibodies that neutralise Herpes simplex virus (HSV)
XX types 1 and 2 - used for diagnosis and passive immuno-therapy of HSV
XX infections.
XX
XX Example 1; Page 81; 100pp; English.
XX
XX AA094730 and AA094731 are PCR primers used to amplify the cpIII membrane
XX anchor gene which is used in part of a process which produces a plasmid
XX pComb-III'. pCombIII' is a bacteriophage vector used to identify human
XX monoclonal antibodies that immunoreact with HSV. One such antibody is the
XX human anti-herpes monoclonal antibody clone FabHSV8. This antibody is
XX capable of neutralising both herpes simplex virus (HSV) types I and II by
XX binding an epitope present on glycoprotein D. The antibody may be used
XX for detecting HSV in vivo or in vitro; for passive immuno-therapy (pref.
XX prophylactically) of HSV infection (eg. genital, oral or ocular herpes),
XX partic. as its Fab fragment and as a competitive reagent for detecting
XX neutralising anti-HSV antibodies in a sample. Anti-Idiotypic antibodies
XX raised against the mAb can be used for active immuno-therapy of HSV
XX infection. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;
Query Match 4.3%; Score 18; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14
RESULT 6
AA0878/c
ID AA0878 standard; DNA; 48 BP.
XX
XX AA0878;
XX
XX 28-JAN-1997 (first entry)
XX
XX pComb2-3 construction primer, G-3(F).
XX
XX Heavy chain; light chain; variable region; VH; monoclonal antibody; MAb;
XX HIV; human immunodeficiency virus; glycoprotein; gp120; clone;
XX virus infectivity assay; precursor gp160; immunocompetence; VL; human;
XX anti-HIV antibody; detection; HIV infection; ss.
XX
XX Synthetic.
XX
XX WO9602273-A1.
XX
XX 01-FEB-1996.
XX
XX 11-JUL-1995; 95WO-US008743.
XX
XX 18-JUL-1994; 94US-00276852.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Burton DR, Barbas CF, Lerner RA;
XX
XX WPI; 1996-179601/18.
XX
XX Monoclonal antibody binding to VI/V2 loop of HIV gp120 - used in passive
XX immuno-therapy and detection of HIV infection.
XX
XX
```


XX Example; Page 76; 366pp; English.

CC The sequences given in AAT40878-82 are primers which were used in the
 CC construction of the combinatorial vector, pComb2-3. pComb2-3 was
 CC constructed using sequences encoding bacteriophage coat protein, gpII,
 CC membrane anchor domain containing a lacZ promoter region sequence 3' to
 CC the membrane anchor for expression of an antibody light chain, and SpeI
 CC and EcoRI cohesive termini prepared from M13mp18. The use of the lac-B
 CC primer, rather than lac-B, results in an amplification product lacking 29
 CC nucleotides on the 3' end but which is functionally equivalent to the
 CC longer fragment. Further amplification reactions resulted in a PCR DNA
 CC fragment consisting of a 5' SpeI site, a cIII DNA membrane anchor domain
 CC beginning at amino acid residue 198 of the mature cIII protein, an
 CC endogenous stop codon provided by the membrane anchor at amino acid 112,
 CC a NheI restriction site, a lacZ promoter, operator and Cap binding site,
 CC and a 3' EcoRI site. Insertion of the coding sequences for antibody light
 CC and heavy chains resulted in the expression of a heavy chain-cIII fusion
 CC product and the light chain protein. The expressed proteins are directed
 CC to the periplasmic space by the pelB leader sequence

SO Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;

Query Match 4.3%; Score 18; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

67 ATGAGAGCCACGCCAC 84
 |||||
 31 ATGAGAGCCACGCCAC 14

RESULT 7
 AAT58272/c
 ID AAT58272 standard; DNA; 48 BP.

AC AAT58272;
 DT 22-AUG-1997 (first entry)

DE M13mp18 PCR primer involved in the expression of fusion proteins.

XX Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;
 KW pharmaceutical; health care; skin treatment; pesticide; herbicide;
 KM heavy metal; polymerase chain reaction; ss.

XX Synthetic.

XX MO9639518-A1.

XX 12-DEC-1996.

XX 05-JUN-1996; 96WO-US009258.

XX 05-JUN-1995; 95US-00462798.

XX 10-OCT-1995; 95US-00541373.

XX (BION-) BIONBRASKA INC.

XX Wylie DE, Lopez O, Murray PJ;

XX WPI; 1997-043140/04.

XX DNA encoding heavy metal binding polypeptide sequences - used for
 PT detecting; removing; adding or neutralising heavy metals, such as lead
 PT cations.

XX Example 3; Page 41; 125pp; English.

XX The present sequence represents the PCR primer for the amplification of
 CC M13mp18 which is involved in the expression of fusion proteins including
 CC heavy chain Fd fragments and light chains from lead cation monoclonal
 CC antibodies. The monoclonal antibody heavy and light chain sequences are

CC derived from RNA isolated from hybridoma cells from mouse spleen cells.
 CC The proteins produced can be used for binding heavy metals, such as lead
 CC cations. They can be used for detecting, removing, adding or neutralising
 CC the heavy metals in biological and inanimate systems. They can be used in
 CC e.g. aqueous liquid systems, in biological or environmental systems or in
 CC such compositions as perfumes, cosmetics, pharmaceuticals, health care
 CC products, skin treatment products, pesticides, herbicides, solvents used
 CC in the production of semi-conductor and integrated circuit components and
 CC production materials for electronic components. The products can provide
 CC for applications involving minute amounts of specific heavy metals

SO Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;

Query Match 4.3%; Score 18; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

67 ATGAGAGCCACGCCAC 84
 |||||
 31 ATGAGAGCCACGCCAC 14

RESULT 8
 AAA32124/c
 ID AAA32124 standard; DNA; 48 BP.

AC AAA32124;

DT 04-JUL-2000 (first entry)

DE Primer G-3 F used in the preparation of pCombIII.

XX Antibody; anti-HIV monoclonal antibody; glycoprotein-120;
 KW human immunodeficiency virus type 1; HIV-1; infectivity titre;
 KM passive immunotherapy; reduce severity; HIV-induced disease;
 KW immunocompetence; active immunisation; ss.

XX Synthetic.

XX AU9948754-A.

XX 17-FEB-2000.

XX 16-SEP-1999; 99AU-00048754.

XX 16-SEP-1999; 99AU-00048754.

XX (SCRI) SCRIPPS RBS INST.

XX Burton DR, Barbas CF, Lerner RA;

XX WPI; 2000-246867/22.

XX Human neutralizing monoclonal antibodies to human immunodeficiency virus
 PT (HIV) used for providing passive immunotherapy to HIV are specific for
 PT glycoprotein-120.

XX Example 1; Page 76; 374pp; English.

XX This sequence represents a polynucleotide used in the preparation of the
 CC antibodies of the invention. The invention relates to the production of
 CC an anti-HIV (human immunodeficiency virus) glycoprotein (gp)-120
 CC monoclonal antibody capable of reducing an HIV infectivity titre in an in
 CC vitro virus infectivity assay by 50% at a concentration of less than 70
 CC ng/ml. The method for the production of the antibody comprises: (a)
 CC providing a first polynucleotide encoding a heavy chain immunoglobulin
 CC amino acid sequence (which does not comprise the sequence represented by
 CC AA99206) and a second polynucleotide encoding a light chain
 CC immunoglobulin amino acid sequence; (b) inserting the first and second
 CC polynucleotide sequences into a host cell; (c) maintaining the host cell
 CC in conditions which allow the amino acid sequences encoded by the
 CC polynucleotides to be expressed in the host cell; and (d) isolating the
 CC antibody comprising the heavy and light chain immunoglobulin amino acid

CC sequences from the host cell. The anti-HIV gp-120 monoclonal antibody is
CC used for providing passive immunotherapy to HIV in a human. They can be
CC administered to high-risk patients to reduce the likelihood and/or
CC severity of HIV-induced disease and to patients who are already HIV-
CC infected. The antibodies are used for neutralising field isolates which
CC provide information about the immunocompetence of an immune response in
CC HIV patients, for detecting HIV in a biological fluid or tissue sample
CC e.g. by radioimmunoassay, for producing anti-idiotypic antibodies which
CC can be used for active immunisation and to screen human monoclonal
CC antibodies to identify those with the same binding specificity and to
CC monitor the course of HIV disease therapy by measuring the changes in
CC concentration of HIV present in the body or in body fluids by
CC immunosay. The anti-HIV gp-120 monoclonal antibodies are encoded by a
CC human polynucleotide sequence and when used in vivo for diagnosis and
CC immunotherapy of HIV-induced disease reduce the problems of significant
CC host immune response to the antibodies associated with monoclonal
CC antibodies of xenogeneic or chimeric derivation

SO Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;

Query Match 4.3%; Score 18; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 9

AAA30998/c
ID AAA30998 standard; DNA; 48 BP.

AC AAA30998;

DT 30-JUN-2000 (first entry)

DE Primer G-3 F used in the preparation of pCOMBIII.

XX Anti-human immunodeficiency virus type 1 antibody; HIV-1; neutralise;
KW reduce HIV infection; diagnosis; immunotherapy; HIV induced disease; ss;
KM glycoprotein 120; gp120; glycoprotein 41; gp41; monoclonal antibody.

XX Synthetic.

PM AU9948756-A.

XX 17-FEB-2000.

PF 16-SEP-1999; 99AU-00048756.

PR 16-SEP-1999; 99AU-00048756.

XX (SCRI) SCRIPPS RES INST.

PI Burton DR, Barbas CF, Lerner RA;

DR WPI; 2000-293393/26.

XX Novel human monoclonal antibodies which immunoreact with and neutralize
PT human immunodeficiency virus useful for treating HIV infections.

PS Example 1; Page 76; 366pp; English.

XX The present sequence is used in the production of anti-human
CC immunodeficiency virus type 1 (HIV-1) antibodies. The invention relates
CC to a human whole immunoglobulin (Ig) molecule which immunoreacts with HIV
CC mature glycoprotein gp120 preferentially over HIV precursor glycoprotein
CC gp160 and neutralises HIV and which reduces HIV infectivity titre in an
CC in vitro virus infectivity assay by 50%, at a concentration of less than
CC 700 ng/ml. The antibodies are used as reagents for the diagnosis and
CC immunotherapy of HIV induced disease. They are useful as neutralising
CC field isolates and provide useful information regarding the

CC immunocompetence of an immune response in HIV infected patients. The
CC monoclonal antibodies are useful for producing anti-idiotypic antibodies
CC which can be used to screen human monoclonal antibodies to identify
CC whether the antibody has the same binding specificity as the antibodies
CC of the invention. The neutralising antibodies define new epitopes on the
CC HIV gp120 and gp41 glycoproteins, thus increasing the availability of new
CC immunotherapeutic human monoclonal antibodies. A major advantage of the
CC monoclonal antibodies derives from the fact that they are encoded by a
CC human polynucleotide sequence. Thus in vivo use of the monoclonal
CC antibodies for diagnosis and immunotherapy of HIV induced disease greatly
CC reduces the problems of significant host immune response to the passively
CC administered antibodies which is a problem commonly encountered when
CC monoclonal antibodies of xenogeneic or chimeric derivation are utilized.
CC An additional major advantage of the monoclonal antibodies described
CC derives from the fact that they immunoreact with a unique determinant
CC present on mature HIV glycoprotein gp120. This class of antibodies is
CC particularly effective at neutralising field isolates of HIV

SO Sequence 48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;

Query Match 4.3%; Score 18; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 10

AAL38725/c
ID AAL38725 standard; DNA; 48 BP.

AC AAL38725;

DT 02-SEP-2002 (first entry)

DE gpIII membrane anchor gene PCR primer SEQ ID No 19.

XX Antibody; epitope; solid support; phage display combinatorial antibody;
KW immunocomplex; antigen; PCR; primer; ss.

XX Unidentified.

PM US6376170-B1.

XX 23-APR-2002.

PF 18-NOV-1997; 97US-00972564.

PR 03-OCT-1994; 94US-00316914.

XX (SCRI) SCRIPPS RES INST.

PI Burton DR, Burioni R, Williamson RA, Sanna PP;

DR WPI; 2002-478279/51.

XX Obtaining antibody that binds to previously unknown epitope on antigen
PT using solid phase capture complex of known antibody-antigen, is useful to
PT screen e.g. phage display combinatorial libraries.

PS Example 1; Col 22; 27pp; English.

XX The invention relates to a novel method for obtaining an antibody that
CC binds to a previously unknown epitope on a preselected antigen, which
CC comprises contacting the preselected antigen with a specific antibody
CC bound to a solid support, contacting the formed immunocomplex with a
CC combinatorial library of antibodies so that a second antibody binds to a
CC previously unknown epitope, and removing the second antibody. The method
CC is used to screen libraries of cloned antibodies, such as phage display
CC combinatorial antibodies. This polynucleotide sequence represents a PCR
CC primer relating to the invention

XX	
Sequence	48 BP; 10 A; 9 C; 15 G; 14 T; 0 U; 0 Other;
5Q	

Query Match	4.3%;	Score 18;	DB 6;	Length 48;
Best Local Similarity	100.0%;	Pred. No. 1e+02;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 11
ADB84272/c
ID ADB84272 standard; DNA; 48 BP

Query Match	4.3%	Score 18;	DB 8;	Length 48;
Best Local Similarity	100.0%	Pred.No. 1e+02;		
Matches 18; Conservative	0;	Mismatches	0;	Gaps 0

QY 67 ATGGAGGCCACCGCCAC 84

Db 31 ATGAGAGCCACCGCCAC 14

RESULT 12
ABN38420
ID ABN38420 standard; DNA; 60 BP.

Query Match	4.1%;	Score 17;	DB 6;	Length 60;
Best Local Similarity	100.0%;	Pred. No. 3.2e+02;		
Matches 17;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

390 GAACCTCAGGCTCCTGG 406

DB 37 GAACCTCAGCTCTCTGG 53

RESULT 13

AAI18458 standard; DNA; 66 BP.

AAI18458;

13-MAY-1999 (first entry)

PCR primer tau3-1 used in construction of hIFNtau3C coding sequence.

Interferon tau3; human, hIFNtau3C; antitumour agent; autoimmune disease; antiviral activity; therapy; PCR primer; ss.

Synthetic.

Homo sapiens.

JP11042089-A.

16-FEB-1999.

29-JUL-1997; 97JP-00203137.

29-JUL-1997; 97JP-00203137.

(SANY) SANKYO CO LTD.

WPI; 1999-197821/17.

An interferon tau 3 modified body protein and polynucleotide - useful for treatment of autoimmune diseases, and in antitumor agents.

Disclosure; Page 6; 18pp; Japanese.

This sequence represents a primer used in the construction of DNA encoding the modified human interferon tau 3 protein of the invention, designated hIFNtau3C. hIFNtau3C is useful as an antitumor agent, and for treatment of autoimmune diseases. The interferon tau 3 modified body hIFNtau3C has a high antiviral activity and is low in cytotoxicity

Sequence 66 BP; 14 A; 19 C; 19 G; 14 T; 0 U; 0 Other;

Query Match 4.1%; Score 17; DB 2; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

390 GAACCTCAGCTCTCTGG 406

42 GAACCTCAGCTCTCTGG 58

RESULT 14

ABZ61630 standard; RNA; 17 BP.

ABZ61630;

21-MAR-2003 (first entry)

Human H-Ras DNAzyme target #421.

Human; ribozyme; short interfering RNA; siRNA; HER2; K-Ras; enzymatic nucleic acid; H-Ras; N-Ras; HIV; cytostatic; anti-HIV; anti-rheumatic; cancer; AIDS; ss.

Homo sapiens.

WO200297114-A2.

05-DEC-2002.

PF 29-MAY-2002; 2002WO-US016840.

XX 29-MAY-2001; 2001US-0294140P.

PR 06-JUN-2001; 2001US-0296249P.

PR 10-SEP-2001; 2001US-0318471P.

XX (RIBO-) RIBOZYME PHARM INC.

PI Mcswiggen U;

XX WPI; 2003-140484/13.

Novel short interfering RNA and enzymatic nucleic acid useful for treating cancer, modulates the expression of a nucleic acid encoding HER2, K-Ras, H-Ras, N-Ras, and human deficiency virus sequences.

Claim 58; Page 119; 185pp; English.

The invention relates to a novel short interfering RNA (siRNA) nucleic acid molecule or an enzymatic nucleic acid molecule, that modulates expression of a nucleic acid molecule encoding HER2, K-Ras, H-Ras, N-Ras, human immunodeficiency virus (HIV) or a component of HIV. The nucleic acid molecule of the invention has cytostatic, anti-HIV, and anti-rheumatic activity. The nucleic acid molecules are useful for reducing HER2, K-Ras, H-Ras, and HIV activity in a cell. The nucleic acids are also useful for treating breast, ovarian, colorectal, lung, prostate, bladder, or pancreatic cancer, and HIV infection, and AIDS. The sequences shown in ABZ59889 - ABZ62216, ABZ64544 - ABZ65531, ABZ6520 - ABZ6524, ABZ6530 - ABZ6585 represent substrate/target sequences for the human ribozymes of the invention

Sequence 17 BP; 4 A; 6 C; 6 G; 0 T; 1 U; 0 Other;

Query Match 3.6%; Score 15; DB 7; Length 17;

Best Local Similarity 93.3%; Pred. No. 3.3e+03;

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

194 ACAGCGCCATCGCGG 208

1 ACAGCGCCATCGCGG 15

RESULT 15

AAV85968/c

AAV85968 standard; DNA; 20 BP.

AAV85968;

10-FEB-1999 (first entry)

Mouse LRP-3 cDNA PCR primer 414r (MUCD 7R).

LRP5; LDL-receptor related protein; LRP-3; IDDM; diagnosis; endocytosis; insulin dependent diabetes mellitus; autoimmune disease; glomerulonephritis; inflammation; viral infection; osteoporosis; hypercholesterolemia; Alzheimer's disease; low density lipoprotein; PCR primer; ss.

Synthetic.

Mus sp.

WO9846743-A1.

22-OCT-1998.

15-APR-1998; 98WO-GB001102.

15-APR-1997; 97US-0043553P.

05-JUN-1997; 97US-0048740P.

(WELL) WELLCOME TRUST LTD.

(MERI) MERCK & CO INC.

PI Todd JA, Hess JM, Caskey CT, Cox RD, Gerhold D, Hammond H;
PI Hey P, Kawaguchi Y, Merriman TR, Metzker ML, Nakagawa Y;
PI Phillips MS, Twells RCJ;

DR WPI; 1998-594573/50.

XX New isolated LDL-receptor related protein - used to develop products for
PT treating, e.g. elevated triglyceride levels, diabetes, autoimmune
PT disorders, inflammation or Alzheimer's disease.

PS Claim 12; Page 117; 2000p; English.

XX The present invention describes LRPS (low density lipoprotein (LDL)
CC receptor related protein, previously designated LRP-3). Nucleic acid
CC molecules (NMs) encoding LRPs can be used for determining if an
CC individual is susceptible to insulin dependent diabetes mellitus (IDDM).
CC The NMs or proteins can be used for reducing triglyceride levels in the
CC serum of an individual. Therapies that affect LRPS may also be useful in
CC the treatment of autoimmune diseases such as glomerulonephritis, diseases
CC and disorders involving disruption of endocytosis and/or antigen
CC presentation, cytokine clearance and/or inflammation, viral infection,
CC pathogenic bacterial toxin contamination, elevation of free fatty acids
CC or hypercholesterolemia, type 2 diabetes, osteoporosis, Alzheimer's
CC disease and cardiovascular disease. Products from the present invention
CC can also be used for detection, diagnosis and drug screening. AAV85917 to
CC AAV86012 represent PCR primers for obtaining LRP-3 cDNA

XX Sequence 20 BP; 3 A; 7 C; 4 G; 6 T; 0 U; 0 Other;

Query Match 3.6%; Score 15; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 3.3e+03;

-Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 GGAGTGTCTGCACAG 179

DB 20 GGAGTGTCTGCACAG 6

Search completed: March 22, 2004, 03:30:38
Job time : 395 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 23:24:50 ; Search time 2999 Seconds

(without alignments)
6026.691 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417
Sequence: 1 atggg99gaactgtgcgcag.....ggtctctgttaccacgarn 417

Scoring table: OLIGO_NUC

Gapop 60.0, Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl.*
1: gb_ba.*
2: gb_bhg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ov.*
8: gb_ov.*
9: gb_ov.*
10: gb_ov.*
11: gb_ov.*
12: gb_ov.*
13: gb_ov.*
14: gb_ov.*
15: gb_ov.*
16: gb_ov.*
17: gb_ov.*
18: gb_ov.*
19: gb_ov.*
20: gb_ov.*
21: gb_ov.*
22: gb_ov.*
23: gb_ov.*
24: gb_ov.*
25: gb_ov.*
26: gb_ov.*
27: gb_ov.*
28: gb_ov.*
29: gb_ov.*
30: gb_ov.*
31: gb_ov.*
32: gb_ov.*
33: gb_ov.*
34: gb_ov.*
35: gb_ov.*
36: gb_ov.*
37: gb_ov.*
38: gb_ov.*
39: gb_ov.*
40: gb_ov.*
41: gb_ov.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	351	84.2	2621	9	BC018142	BC018142 Homo sapi
2	351	84.2	3682	9	AY032927	AY032927 Homo sapi
3	351	84.2	3766	6	AX746790	AX746790 Sequence
4	351	84.2	3766	9	AK091123	AK091123 Homo sapi
5	351	84.2	3931	9	AF328642	AF328642 Homo sapi
6	212	50.8	55173	2	AC132816	AC132816 Homo sapi
7	212	50.8	144000	9	AC123764	AC123764 Homo sapi
8	212	50.8	182016	9	AC087741	AC087741 Homo sapi
9	212	50.8	187865	2	AC015559	AC015559 Homo sapi
10	212	50.8	207822	2	AC109321	AC109321 Homo sapi
11	205	49.2	55173	2	AC132816	AC132816 Homo sapi
12	161	38.6	74301	2	AC137736	AC137736 Homo sapi
13	33	7.9	3772	10	AF363457	AF363457 Mus muscu
14	33	7.9	3995	10	BC029102	BC029102 Mus muscu
15	33	7.9	244653	10	AL645911	AL645911 Mouse DNA
16	23	5.5	3164	10	BC004692	BC004692 Mus muscu
17	22	5.5	160170	2	AP001158	AP001158 Homo sapi
18	22	5.3	221941	9	AC091060	AC091060 Homo sapi
19	22	5.3	224788	2	AP001905	AP001905 Homo sapi
20	21	5.0	1260	10	AB010281	AB010281 Mus muscu
21	21	5.0	4085	10	AY135367	AY135367 Mus muscu
22	21	5.0	65608	2	AC132197	AC132197 Homo sapi
23	21	5.0	110000	2	LMFLCH36-24	Continuation (25 o
24	21	5.0	165614	10	AC111044	AC111044 Mus muscu
25	21	5.0	182626	2	AC124374	AC124374 Mus muscu
26	21	5.0	225027	2	AC131725	AC131725 Mus muscu
27	21	5.0	239392	2	AC115417	AC115417 Rattus no
28	21	5.0	245032	2	AC109749	AC109749 Rattus no
29	20	4.8	45013	9	AL138684	AL138684 Human DNA
30	20	4.8	53730	9	AC135726	AC135726 Homo sapi
31	20	4.8	135351	9	AC068279	AC068279 Homo sapi
32	20	4.8	150489	2	AC068854	AC068854 Homo sapi
33	20	4.8	165028	9	AC110758	AC110758 Homo sapi
34	20	4.8	183176	2	AC106894	AC106894 Homo sapi
35	20	4.8	191947	2	AC135875	AC135875 Rattus no
36	20	4.8	192181	9	AC133644	AC133644 Homo sapi
37	20	4.8	201764	2	AL359818	AL359818 Homo sapi
38	20	4.8	202312	2	AC135269	AC135269 Rattus no
39	20	4.8	204006	2	AC134881	AC134881 Homo sapi
40	20	4.8	205662	9	AC017002	AC017002 Homo sapi
41	20	4.8	214361	2	AC105814	AC105814 Rattus no
42	20	4.8	229276	2	AC120704	AC120704 Rattus no
43	20	4.8	240214	2	AC126871	AC126871 Rattus no
44	20	4.8	283969	2	AC120737	AC120737 Rattus no
45	19	4.6	2294	10	BC004088	BC004088 Mus muscu

ALIGNMENTS

RESULT 1
BC018142 2621 bp mRNA linear PRI 03-OCT-2003
LOCUS Homo sapiens caspase recruitment domain family, member 14, mRNA
DEFINITION (CDNA clone MGC:9539 IMAGE:3847282), complete cds.
ACCESSION BC018142
VERSION BC018142.1 GI:17390314
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2621)
Strausberg,R.L., Feingold,E.A., Grouse,J.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT												
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.U., Uedini, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulhly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, T.A., Gunzette, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultx, S.W., Villalon, D.K., Muzny, D.M., Sodegren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Hailton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schercher, A., Schein, J.E., Jones, S.J., and Marra, M.A.	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)	22388257	12477932	2 (bases 1 to 2621)	Strausberg, R.	Direct Submission	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA	NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk Email: cgapds-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.sngc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.												
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAX Library: 20 Row: 1 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16507953.	Location/Qualifiers	1. .2621	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="MGC:9539 IMAGE:1847282"	/tissue_type="Colon, adenocarcinoma"	/clone_id="N1H_MGC_65"	/lab_host="DH103"	/note="vector: pCMV-SPORT6"	1. .2621	/gene="CARD14"	/note="synonyms: CARMA2, BIMP2"	/db_xref="LocusID:79092"	/db_xref="MIM:607211"	178. .2400	/codon_start=1	/product="CARD14 protein"	/protein_id="AA18142.1"	/db_xref="GI:17390315"	/translation="MSELCDSDALTLDEETLWEMSHRRIVRCIPSLTPYLA OAKYLTQDEEVHSPRLTNSARAGHLLTLKRGNGAIAFLSKFNPVYTL VTGIGPVDPSVPSGLMETSKLTFCIGAGLSLOELNIOEGCKEVLIRPCOIOEH GLATRAEGHQLADHSRKREYSAFHEVTLRLKDEMLSLSTASNLOKEKLAAS CRSLQDELVTLLKQELQANWVSSELELQEOSLTASDQSGDELNLRKENEKLR LTFIAKDLITQSLIDKRAVGRSRLVETIHILRRRAVAERQDQWVEKQTLIQF

misc_feature	1150..1404
/note="Myosin_tail; Region: Myosin tail. The myosin molecule is a multi-subunit complex made up of two heavy chains and four light chains it is a fundamental contractile protein found in all eukaryote cell types. This family consists of the coiled-coil myosin heavy chain tail region. The coiled-coil is composed of the tail from two molecules of myosin. These can then assemble into the macromolecular thick filament. The coiled-coil region provides the structural backbone the thick filament"	1924..2142
/db_xref="CD0:pfam01576"	
/note="PDZ; Region: Domain present in PSD-95, Dlg, and ZO-1/2. Also called DHR (Dlg homologous region) or GGF (relatively well conserved tetrapeptide in these domains). Some PDZs have been shown to bind C-terminal polypeptides"	
/db_xref="CD0:smart00228"	
Query Match	84.2%; Score 351; DB 9; Length 2621;
Best Local Similarity	100.0%; Pred. No. 2,8e-170;
Matches 351; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGGGGGAAGCTGTGCGCGCAGAGGACTCGCACTCAGCGCACTGAGCAGAGAGACACTGTGG 60
DB	178 ATGGGGGAAGCTGTGCGCGCAGAGGACTCGCACTCAGCGCACTGAGCAGAGAGACACTGTGG 237
QY	61 GAGATGATGGAGAGCCACCGGCACAAGGATGTAGCGCTGACCCAGCGGCTCAAC 120
DB	238 GAGATGATGGAGAGCCACCGGCACAAGGATGTAGCGCTGACCTGCCAGCGGCTCAAC 297
QY	121 CCTTACCTGTGGCAGAGCCAGAGTGTGTGTGCAGCTGAGCAGAGAGAGAGTGTCTTCAACGC 180
DB	298 CCTTACTGTGGCAGAGCCAGAGTGTGTGTGCAGCTGAGCAGAGAGAGAGTGTCTTCAACGC 357
QY	181 CCCCGGCTCAACCAACAGCGGCATATGCGGCGCGGCACATCTTGCTGATTTGCTGAAGACTCGA 240
DB	358 CCCCGGCTCAACCAACAGCGGCATATGCGGCGCGGCACATCTTGCTGATTTGCTGAAGACTCGA 417
QY	241 GGGAGAAACGGGGGCATATGCGGCTTCTGTGAGAGAGCTGAAGTCCACCAACCTGAGAGCTTAC 300
DB	418 GGGAGAAACGGGGGCATATGCGGCTTCTGTGAGAGAGCTGAAGTCCACCAACCTGAGAGCTTAC 477
QY	301 ACCCTGTGTCAACGGGGCTGACAGCTGATGTGACTTTCAGTAACTTTAGCGGT 351
DB	478 ACCCTGTGTCAACGGGGCTGACAGCTGATGTGACTTTCAGTAACTTTAGCGGT 528
RESULT 2	
AY032927	3682 bp mRNA linear PRI 07-DEC-2001
LOCUS	
DEFINITION	Homo sapiens CARD-containing MAGUK 2 protein (CARDMA2) mRNA, complete cds.
ACCESSION	AY032927
VERSION	AY032927.1
KEYWORDS	GI:14192722
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3682)
TITLE	Galde,O., Martinon,F., Micheau,O., Bonnet,D., Thome,M. and Techopp,U. CARD-containing binding partner of Bcl10, induces Bcl10 phosphorylation and NF-kappab activation

JOURNAL	FEBS Lett. 496 (2-3), 121-127 (2001)
MEDLINE	21255663
PUBMED	11356195
REFERENCE	2 (bases 1 to 3682)
AUTHORS	Martinon, F., Thome, M. and Tschopp, J.
TITLE	Direct Substitution
JOURNAL	Submitted (23-APR-2001) Institute of Biochemistry, University of
Lausanne, Cn. des Boveresses 155, Epalinges, VD 1066, Switzerland	
FEATURES	Location/Qualifiers

```

source
1. .3682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene
1. .3682
/gene="CARMA2"
CDS
4. .3018

```

/notes="contains caspase recruitment domain CARD, coiled
 coil", PDZ, SH3, guanylate kinase domain GUK; interacts
 with the CARD-containing molecule Bcl10 (CARDEN)"
 /product="CARD-containing MAGUK 2 protein"
 /codon_start=1
 /protein_id="AAK54453.1"
 /db_xref="tgi:14192723"
 /translation="MGEELCRDQNALALDEETLMMEMSHRRIRYRCSPRLDPTYL
 QAYQLQDDEEYVSHSRILNLSMAEGLAGLIGSLQELTKGKQALVLSLRKHPDPTYL
 QVQLQVDPDSESSNGMLTESLTPKAGLIGSLQELTKGKQALVLSLRKHPDPTYL
 GLATRAEGHLHLEADHSRRKRSVAHFPEVLTAKDMESLSLHNSALQENSLAAS
 CRSLDEELVLTKEOLRQANVNSCELELOEOSTLRTPDSDSEELNLTKEENLR
 LTRSLAEKDLLEQSLDPAKRSRRELYERLHSLREPVLAERQREYKESQTLFQ
 KSRQAQYKVEKNALQVQCELRQKDRDYASRDAQFETLSQIVKDSLRQYFET
 TVDCELTQRLQQLQAPPVYLQKARTTRCPREKQQLVNRNALCPDSDSCYVSS
 TEGSLDLSLDAESRELVDFSRSSAPPEQOGLYKRVAVDEEGEPMFSRGLPE
 DQGLPKARKGHPHLDVLLDPTDLPOLESLSQVPGGLDVGSGMFRSGRLPI
 SQYTMLAFOGDLLEQSLVSGNLTGLFTHRTYQKADQMLRGTQVLDVYAS
 PLFPAVLEETLLEAVGLLELVAVDGMFCCLSKVNTDYSKALLDPAKTAISDSYIT
 VNLMEARKEGLLEQVNEVLAIVDTDMFQCCQGMHARRNLSITMDQLAAHGTIPV
 AQOGLILALDPMQCCVITSPKSSGGPKQVLRVLSYADKRAKSLPDLSTPDQGLP
 EGSSLTGMASCSLTLYVLTLYTSSWGPQALPVLPVRAQGLISELCLQKFKCLAL
 EYLSDEYEVANSQSGDILQSGEYVSCQGLTGTSEBQVLLLEAARQEGGLDPAKLY
 SLSLA
 /MD5="F71VHVSHVNRKQAKGGLQGLTGTSEBQVLLLEAARQEGGLDPAKLYSLSLA
 GMSDLDGLLSVCRQALADQKQKVVWTEQSP"

ORIGIN

Query Match	84.2%;	Score 351;	DB 9;	Length 3682;
Best Local Similarity	100.0%;	Pred. No. 2.7e-170;		
Matches 351;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	ATGGGGGAATGTGGCCGCAAGGACTCCGACCTCAAGGCACTGGACGAGAAACAATTGG	60
Db	ATGGGGGAATGTGGCCGCAAGGACTCCGACCTCAAGGCACTGGACGAGAAACAATTGG	63
Qy	GAGATGATGAGAGCCACCGCCACAGGATGTAAGCTGATCTGGCCAGCGCTCAAC	120
Db	GAGATGATGAGAGCCACCGCCACAGGATGTAAGCTGATCTGGCCAGCGCTCAAC	122
Qy	CCCTACCTGCGCCAGGCCAAGTGCTGTGSCCAAGCTGGAAGAGAGAGTGCTGCAAGC	180
Db	CCCTACCTGCGCCAGGCCAAGTGCTGTGSCCAAGCTGGAAGAGAGAGTGCTGCAAGC	183
Qy	CCCCGCTCAACCAACAGCGCCATGCGGGCCGGGCACTTGTGTGAATTTGCTGAAGACTGGA	240
Db	CCCCGCTCAACCAACAGCGCCATGCGGGCCGGGCACTTGTGTGAATTTGCTGAAGACTGGA	243
Qy	GGGAAGAACGGGGCCATTCGCTTCCCTGGAGAGCTGGAAGTTCCACAACCCGAGGCTAC	300
Db	GGGAAGAACGGGGCCATTCGCTTCCCTGGAGAGCTGGAAGTTCCACAACCCGAGGCTTAC	303
Qy	ACCGTGCTCACGGGGCTGACGCTGAATGTTGACTTCAGTAATTTCAGCGGT	351
Db	ACCGTGCTCACGGGGCTGACGCTGAATGTTGACTTCAGTAATTTCAGCGGT	354

RESULT 3				
AX746790	AX746790	3766 bp	mRNA	linear
LOCUS	AX746790			
DEFINITION	Sequence 315 from Patent EP1308459.			
ACCESSION	AX746790			
VERSION	AX746790.1	GI:32131178		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM				

REFERENCE

1. Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yano, T., Imano, Y., Otsuka, K., Nara, K., Irie, R.

TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1308459-A 315 07-MAY-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
Identification/Classification

```

FEATURES
source
Location/Qualifiers
1. .3766

```

ORIGIN

Query Match	84.2%;	Score 351;	DB 6;	Length 3766;
Best Local Similarity	100.0%;	Pred. No. 2.7e-170;		
Matches 351;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

1 ATGGGGAACTGTGCCGAGGACTCCGCACTCACGGCACTGGACGAGGACACTGTGG 60

Db 192 ATGGGGAACTGTGCCCGCAGGACTCCGCACTACCGCACTGGACGAGAGACACTGTGG 251

61 GAGATGATGAGAGCCACCGCCACAGATCGTACGCTGCATCTGCCCCAGCCGCTCACC 120

D_b 252 GAGATGATGAGAGCCACCGCCACAGGATCGTACGCTGCATCTGCCCCAGCCGCTCACC 311

121 CCCTACCTGCGCCAGGCCAAGGTGCTGTGCCAGCTGGACGAGGAGGTGCTGCACAGC 180

312 CCTACCTGCGCCAGGCCAAGTCTGTGCCAGCTGGACGAGGAGGTGCTGCACAGC 371

181 TTTTCTGGATTGATTTTGCTGAAGAAGACCTCCGA 240

431

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

0

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

[illegible]

RESULT 4

AK091123
tOCTIS
AK091123
3766 bp
mRNA
linear
PRI 15-JUL-200

DEFINITION Homo sapiens cDNA FLJ33804 f1s, clone C10NG2000766, weakly similar to Bat1p, a putative caspase recruitment domain protein 9 mRNA.

ACCESSION AK091123
 NUCLEOTIDE SEQUENCE 1 CT:31748419

KEYWORDS oligo capping; fis (full insert sequence).

ORGANISM	Homo sapiens	Canis lupus familiaris	Escherichia coli	Arabidopsis thaliana	Drosophila melanogaster	Mus musculus	Rattus norvegicus	Gallus gallus	Xenopus laevis	Plasma	Yeast	Plant	Fungi	Bacteria	Virus
Genome Size (Mb)	3,000	2,800	4.6	113	180	2,700	2,700	1,000	3,500	0.001	12	113	10	4.6	0.001
Genome Type	Diploid	Diploid	Prokaryotic	Diploid	Diploid	Diploid	Diploid	Diploid	Diploid	Prokaryotic	Eukaryotic	Eukaryotic	Eukaryotic	Prokaryotic	Prokaryotic
Genome Complexity	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Organization	Linear	Linear	Circular	Linear	Linear	Linear	Linear	Linear	Linear	Circular	Linear	Linear	Linear	Circular	Circular
Genome Annotation	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Assembly	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Analysis	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Function	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Evolution	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Conservation	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Variation	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Stability	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Replication	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Transcription	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Translation	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Regulation	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Interaction	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Structure	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Dynamics	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Phenotype	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Environment	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Health	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Disease	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Therapy	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Prognosis	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Prevention	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Cure	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Research	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Education	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Awareness	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Understanding	High	High	Low	Low	Low	High	High	High	High	Low	High	High	High	Low	Low
Genome Knowledge	High	High	Low	Low	Low	High	High	High	High	Low	High	High			

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,

Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsumawa, H., Takahashi, E.

Katoaka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S.,
 Terashima,Y., Matanabe,M., Sugiyama,T., Irie,R., Otsuki,T.,
 Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
 Wagatsuma,M., Murakawa,K., Kanenori,K., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahara,K., Masuo,Y., Nagai,K. and Isega,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3766)
 Isega,T. and Yamamoto,J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isega, FMJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genom@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3966)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology (RAB); cDNA library
 construction; Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing; HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1..3766
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CTONG2000766"
 /issue_type="tongue, tumor tissue"
 /clone_lib="CTONG2"
 /note="Cloning vector: pME18SFL3"
 ORIGIN
 Query Match 84.2%; Score 351; DB 9; Length 3766;
 Best Local Similarity 100.0%; Pred. No. 2.7e-170; Indels 0; Gaps 0;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGGGAAGTCTGTCGGGAGGAGTCCGCACTCAGGCACTGACGAGAGACACTGTGG 60
 DB 192 ATGGGGGAAGTCTGTCGGGAGGAGTCCGCACTCAGGCACTGACGAGAGACACTGTGG 251
 QY 61 GAGATGATGAGAGGACCGCCACAGATGCTACGCTGATCTGCCAGCCGCTCACC 120
 DB 252 GAGATGATGAGAGGACCGCCACAGATGCTACGCTGATCTGCCAGCCGCTCACC 311
 QY 121 CCTTACCTGCGCCGCGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 180
 DB 312 CCTTACCTGCGCCGCGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 371
 QY 181 CCGCGGCTCACCACAGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 240
 DB 372 CCGCGGCTCACCACAGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 431
 QY 241 GGAAGAAGAGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 300
 DB 432 GGAAGAAGAGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 491
 QY 301 ACCCTGTGTCACCGGCTGTCAGGCTGATGTTGACTTCACTTAACTTTAGCGGT 351
 DB 492 ACCCTGTGTCACCGGCTGTCAGGCTGATGTTGACTTCACTTAACTTTAGCGGT 542
 RESULT 5
 AF322642 3931 bp mRNA linear PRI 09-APR-2001
 LOCUS Homo sapiens caspase recruitment domain protein 14 mRNA, complete
 DEFINITION cds.
 ACCESSION AF322642
 VERSION AF322642.1 GI:12382774
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 3931)
 AUTHORS Bertin,J., Wang,L., Guo,Y., Jacobson,M.D., Poyet,J.L.,
 Srinivasula,S.M., Merriam,S., Ditzel,P.S. and Alnemri,E.S.
 TITLE CARD11 and CARD14 are novel caspase recruitment domain
 (CARD)/membrane-associated guanylate kinase (MAGUK) family members
 that interact with Bcl10 and activate NF-kappa B
 JOURNAL J. Biol. Chem. 276 (15), 11877-11882 (2001)
 MEDLINE 21192234
 PUBMED 11278692
 REFERENCES 2 (bases 1 to 3931)
 AUTHORS Bertin,J.
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-2000) Neurobiology, Millennium Pharmaceuticals
 Inc., 640 Memorial Drive, Cambridge, MA 02139, USA
 FEATURES
 source
 1..3931
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /translat="MAGUK family member"
 /function="binds to Bcl10 and activates NF-kB"
 /note="CARD14; CARD/MAGUK family member"
 /codon_start=1
 /product="caspase recruitment domain protein 14"
 /protein_id="AA053403.1"
 /db_xref="GI:12382775"
 CDS
 1..3931
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /translat="MAGUK family member"
 /function="binds to Bcl10 and activates NF-kB"
 /note="CARD14; CARD/MAGUK family member"
 /codon_start=1
 /product="caspase recruitment domain protein 14"
 /protein_id="AA053403.1"
 /db_xref="GI:12382775"
 ORIGIN
 Query Match 84.2%; Score 351; DB 9; Length 3931;
 Best Local Similarity 100.0%; Pred. No. 2.7e-170; Indels 0; Gaps 0;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGGGAAGTCTGTCGGGAGGAGTCCGCACTCAGGCACTGACGAGAGACACTGTGG 60
 DB 207 ATGGGGGAAGTCTGTCGGGAGGAGTCCGCACTCAGGCACTGACGAGAGACACTGTGG 266
 QY 61 GAGATGATGAGAGGACCGCCACAGATGCTACGCTGATCTGCCAGCCGCTCACC 120
 DB 267 GAGATGATGAGAGGACCGCCACAGATGCTACGCTGATCTGCCAGCCGCTCACC 326
 QY 121 CCTTACCTGCGCCGCGGAGGAGTGTCTGCTGCACTGACGAGAGAGGAGTGTCTGCAAGC 180

Db 327 CCTAAGTCCGCGCCAGGCGCAAGTGTGTGCGACGTGACGAGAGAGTGTCTGACACG 386
QY 181 CCCCAGCTCAACCAAGCGCCATGCGGCGCGGCACTTGATTTGCTGAAAGACTCGA 240
Db 387 CCCCAGCTCAACCAAGCGCCATGCGGCGCGGCACTTGATTTGCTGAAAGACTCGA 446
QY 241 GGAAGAAAGGCGGCGCATGCGCTTCTGAGAGAGCTGAAAGTTCCAAACCTGACGCTTAC 300
Db 447 GGAAGAAAGGCGGCGCATGCGCTTCTGAGAGAGCTGAAAGTTCCAAACCTGACGCTTAC 506
QY 301 ACCCTGTGACCGGCTGACGCTGATGTTGACTTCACTTAAGCGGT 351
Db 507 ACCCTGTGACCGGCTGACGCTGATGTTGACTTCACTTAAGCGGT 557

RESULT 6
AC132816
LOCUS
DEFINITION
AC132816 55173 bp DNA linear HTG 04-SEP-2002
SEQUENCE SAMPLING.
AC132816
AC132816.1 GI:22711605
HTG: HTGS_PHASE0.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 55173)
Biren, B., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone RP13-593A24
Unpublished
2 (bases 1 to 55173)
Biren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Batra, N., Baatien, V., Bloom, T., Boguslavsky, L., Bouckgeater, B.,
Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,
Karataas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, U., Peterson, K.,
Phunhahang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Rhmann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainout, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (04-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 128179
Center clone name: 593_A_24

* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
693 692: contig of 692 bp in length
793 792: gap of 100 bp
1480 1479: contig of 687 bp in length
1580 1579: gap of 100 bp
2272 2271: contig of 692 bp in length
2372 2371: gap of 100 bp
3057 3056: contig of 685 bp in length
3157 3156: gap of 100 bp
3859 3858: contig of 702 bp in length
3959 3958: gap of 100 bp
4633 4632: contig of 674 bp in length
4733 4732: gap of 100 bp
5419 5418: contig of 686 bp in length
5519 5518: gap of 100 bp
6200 6199: contig of 681 bp in length
6300 6299: gap of 100 bp
6969 6968: contig of 670 bp in length
7070 7069: gap of 100 bp
7779 7778: contig of 710 bp in length
7879 7878: gap of 100 bp
7880 7879: contig of 689 bp in length
8569 8568: gap of 100 bp
8669 8668: contig of 696 bp in length
9365 9364: gap of 100 bp
9465 9464: gap of 100 bp
10171 10170: contig of 707 bp in length
10172 10171: gap of 100 bp
10272 10271: contig of 692 bp in length
10964 10963: gap of 100 bp
11064 11063: contig of 688 bp in length
11752 11751: gap of 100 bp
11852 11851: contig of 696 bp in length
12547 12546: gap of 100 bp
12647 12646: contig of 662 bp in length
13310 13309: gap of 100 bp
13410 13409: contig of 683 bp in length
14099 14098: gap of 100 bp
14199 14198: contig of 671 bp in length
14870 14869: gap of 100 bp
14959 14958: contig of 698 bp in length
14970 14969: gap of 100 bp
15668 15667: contig of 698 bp in length
15678 15677: gap of 100 bp
16479 16478: contig of 712 bp in length
16579 16578: gap of 100 bp
16580 16579: contig of 689 bp in length
17268 17267: gap of 100 bp
17368 17367: contig of 670 bp in length
18039 18038: gap of 100 bp
18138 18137: contig of 691 bp in length
18139 18138: gap of 100 bp
18829 18828: contig of 691 bp in length
18930 18929: gap of 100 bp
19632 19631: contig of 703 bp in length
19633 19632: gap of 100 bp
19733 19732: contig of 700 bp in length
20433 20432: gap of 100 bp
20533 20532: contig of 681 bp in length
21213 21212: contig of 681 bp in length
21214 21213: gap of 100 bp
21988 21987: contig of 675 bp in length
21989 21988: gap of 100 bp
22089 22088: contig of 713 bp in length
22801 22800: gap of 100 bp
22901 22900: gap of 100 bp
22902 22901: contig of 674 bp in length
23575 23574: contig of 674 bp in length
23576 23575: gap of 100 bp
23675 23674: gap of 100 bp
24355 24354: contig of 680 bp in length
24356 24355: gap of 100 bp
24455 24454: gap of 100 bp
25146 25145: contig of 691 bp in length
25246 25245: gap of 100 bp
25247 25246: contig of 701 bp in length
25947 25946: contig of 701 bp in length
26047 26046: gap of 100 bp
26740 26739: contig of 693 bp in length
26741 26740: gap of 100 bp
26840 26839: gap of 100 bp
27543 27542: contig of 703 bp in length
27544 27543: gap of 100 bp

```

* 27644 28338: contig of 695 bp in length
* 28339 28438: gap of 100 bp
* 28439 29137: contig of 699 bp in length
* 29138 29237: gap of 100 bp
* 29238 29924: contig of 687 bp in length
* 29925 30024: gap of 100 bp
* 30025 30712: contig of 688 bp in length
* 30713 30812: gap of 100 bp
* 30813 31504: contig of 692 bp in length
* 31505 31604: gap of 100 bp
* 31605 32322: contig of 718 bp in length
* 32323 32422: gap of 100 bp
* 32423 33104: contig of 682 bp in length
* 33105 33204: gap of 100 bp
* 33205 33899: contig of 695 bp in length
* 33900 33999: gap of 100 bp
* 34000 34690: contig of 691 bp in length
* 34691 34790: gap of 100 bp
* 34791 35481: contig of 691 bp in length
* 35482 35581: gap of 100 bp
* 35582 36277: contig of 696 bp in length
* 36278 36377: gap of 100 bp
* 36378 37072: contig of 695 bp in length
* 37073 37172: gap of 100 bp
* 37173 37856: contig of 684 bp in length
* 37857 37956: gap of 100 bp
* 37957 38652: contig of 696 bp in length
* 38653 38752: gap of 100 bp
* 38753 39440: contig of 688 bp in length
* 39441 39540: gap of 100 bp
* 39541 40218: contig of 678 bp in length
* 40219 40318: gap of 100 bp
* 40319 41015: contig of 697 bp in length
* 41016 41115: gap of 100 bp
* 41116 41820: contig of 705 bp in length
* 41821 41920: gap of 100 bp
* 41921 42600: contig of 680 bp in length
* 42601 42700: gap of 100 bp
* 42701 43386: contig of 686 bp in length
* 43387 43486: gap of 100 bp
* 43487 44179: contig of 693 bp in length
* 44180 44279: gap of 100 bp
* 44280 44939: contig of 660 bp in length
* 44940 45039: gap of 100 bp
* 45040 45703: contig of 664 bp in length
* 45704 45803: gap of 100 bp
* 45804 46470: contig of 667 bp in length
* 46471 46570: gap of 100 bp
* 46572 47264: contig of 694 bp in length
* 47265 47364: gap of 100 bp
* 47365 48069: contig of 705 bp in length
* 48070 48169: gap of 100 bp
* 48170 48847: contig of 678 bp in length
* 48848 48947: gap of 100 bp
* 48948 49636: contig of 689 bp in length
* 49637 49737: gap of 100 bp
* 49738 50437: contig of 701 bp in length
* 50438 50537: gap of 100 bp
* 50538 51213: contig of 676 bp in length
* 51214 51313: gap of 100 bp
* 51314 51999: contig of 686 bp in length
* 52000 52099: gap of 100 bp
* 52100 52792: contig of 693 bp in length
* 52893 52892: gap of 100 bp
* 52893 53594: contig of 702 bp in length
* 53595 53694: gap of 100 bp
* 53695 54377: contig of 683 bp in length
* 54378 54477: gap of 100 bp
* 54478 55173: contig of 696 bp in length.

```

FEATURES
Source Location/Qualifiers
1..55173

Query Match 50.8%; Score 212; DB 2; Length 55173;

```

Best Local Similarity 100.0%; Pred. No. 1.6e-98;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAAGTGTGGCCGAGGAGTCCGACTCAGCGCACTGAGCAGAGAGACTGTGG 60
Db 51604 ATGGGGGAAGTGTGGCCGAGGAGTCCGACTCAGCGCACTGAGCAGAGAGACTGTGG 51663
QY 61 GAGATGATGAGAGAGCAGCGCCAGCAGAGATGCTAGCGTGCATCTGCCAGCGCTCAGC 120
Db 51664 GAGATGATGAGAGAGCAGCGCCAGCAGAGATGCTAGCGTGCATCTGCCAGCGCTCAGC 51723
QY 121 CCTACTCTGGCCGAGGAGCAGAGTGTGTGCGCAGCTGAGCAGAGAGAGTGTCTGACAGC 180
Db 51724 CCTACTCTGGCCGAGGAGCAGAGTGTGTGCGCAGCTGAGCAGAGAGAGTGTCTGACAGC 51783
QY 181 CCCGGGCTCAGCAAGCGCCATGCGGCGCG 212
Db 51784 CCCGGGCTCAGCAAGCGCCATGCGGCGCG 51815

RESULT 7
AC123764
LOCUS Homo sapiens chromosome 17, clone CTC-784N20, complete sequence.
DEFINITION AC123764
ACCESSION AC123764.18 GI:28913166
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 144000)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone CTC-784N20
Unpublished
2 (bases 1 to 144000)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouhgalter,B., Brown,A., Camarata,J., Campiagno,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Govette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacombe,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Margulis,N.,
Mathews,C., McCarthy,M., McEwan,P., McKernan,K., Meltrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicot,R., Norbu,C., Norman,C.H., O'Connor,I., O'Donnell,E.,
O'Neill,D., Oliver,J., Peterson,K., Plunkharg,M., Riley,R., Rise,C.,
Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Teste,S.,
Theodore,J., Topham,K., Travers,M., Travins,N., Triggillo,D.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.,J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 144000)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collimore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Govette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacombe,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Margulis,N.,
Mathews,C., McCarthy,M., McEwan,P., McKernan,K., Meltrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicot,R., Norbu,C., Norman,C.H., O'Connor,I., O'Donnell,E.,
O'Neill,D., Oliver,J., Peterson,K., Plunkharg,M., Riley,R., Rise,C.,
Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupbach,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Teste,S.,
Theodore,J., Topham,K., Travers,M., Travins,N., Triggillo,D.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.,J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

TITLE
JOURNAL
REFERENCE
AUTHORS

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (16-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 144000)

Birren, B., Nusbaum, C., Lander, E., Abouelella, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeAtellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, D., Gargya, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Mathews, C., McCarthy, M., Melrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (26-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

5 (bases 1 to 144000)

Birren, B., Nusbaum, C., Lander, E., Abouelella, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeAtellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gargya, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Mathews, C., McCarthy, M., Melrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (11-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 11, 2003 this sequence version replaced gi127902327.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 127277
Center clone name: 784_N_20

Only the first 144,000 kilobases of this clone are being submitted.

FEATURES
source
The remainder overlaps accession number AC124319 [WICGR project 127275].

Location/Qualifiers
1. 144000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="17"
/clone_17b="CTC-784N20"
repeat_region
212..264
/rpt_family="WIR"
complement(2756..1030)
/rpt_family="L2"
repeat_region
1977..2009
/rpt_family="(TG)n"
2186..2304
/rpt_family="(TCCC)n"
complement(2305..2587)
/rpt_family="Alusx"
2618..2752
/rpt_family="CT-rich"
complement(2756..2929)
/rpt_family="FRAM"
2939..2962
/rpt_family="(TTTGT)n"
complement(4091..4184)
/rpt_family="MER94"
complement(4330..4638)
/rpt_family="Alusx"
4708..4756
/rpt_family="CT-rich"
4760..4782
/rpt_family="(CA)n"
complement(4876..4946)
/rpt_family="HAL1b"
4976..5002
/rpt_family="AT-rich"
complement(5003..5050)
/rpt_family="MER65A"
complement(5058..5349)
/rpt_family="Alusg"
5350..5382
/rpt_family="AT-rich"
5510..5651
/rpt_family="BC200"
complement(5658..5963)
/rpt_family="Alusx"
complement(6417..6730)
/rpt_family="Alub"
complement(7421..7543)
/rpt_family="FLAM C"
complement(7560..7591)
/rpt_family="LIMB8"
7592..7909
/rpt_family="AluJo"
complement(7910..8041)
/rpt_family="LIMB8"
8112..8231
/rpt_family="MER45A"
complement(8406..8661)
/rpt_family="MLTIC"
complement(8662..8856)
/rpt_family="Alub"
8857..9148
/rpt_family="Alusx"
complement(9149..9286)
/rpt_family="Alub"
complement(9287..9441)
/rpt_family="MLTIC"
complement(10183..10488)
/rpt_family="Alusx"


```
repeat_region /rpt_family="MER5B"
6661..6763
/rpt_family="MER5A"
repeat_region complement(6879..7187)
/rpt_family="AluXs"
7241..7336
/rpt_family="MER81"
7332..7514
/rpt_family="MER20"
8710..8732
/rpt_family="(GAAA)n"
repeat_region complement(8788..8868)
/rpt_family="MER20"
9120..9606
/rpt_family="MER1A"
9637..9943
/rpt_family="AluXs"
repeat_region complement(10268..10576)
/rpt_family="AluXs"
repeat_region complement(10582..10738)
/rpt_family="FRAM"
repeat_region complement(10806..10962)
/rpt_family="AluSc"
complement(10958..10973)
/nc="single clone coverage"
11049..11236
/rpt_family="AluDb"
12381..12509
/rpt_family="Char1168"
repeat_region complement(12611..12920)
/rpt_family="AluSg"
repeat_region complement(12949..13219)
/rpt_family="AluSc"
repeat_region complement(15127..15161)
/rpt_family="L2"
complement(15228..15283)
/rpt_family="MIR"
15284..15647
/rpt_family="THEIC"
repeat_region complement(15648..15685)
/rpt_family="MIR"
repeat_region complement(15782..15860)
/rpt_family="L2"
17124..17422
/rpt_family="AluXs"
complement(17967..18101)
/rpt_family="L1ME4A"
complement(19532..19840)
/rpt_family="AluXs"
repeat_region complement(21106..21402)
/rpt_family="AluO"
22639..23145
/rpt_family="L1MC4a"
23146..23203
/rpt_family="(TATATG)n"
repeat_region complement(23323)
/rpt_family="L1MC4a"
23324..23637
/rpt_family="AluXs"
23638..23851
/rpt_family="L1MC4a"
24094..24625
/rpt_family="L1MC4a"
24829..25117
/rpt_family="AluO"
complement(25762..25901)
/rpt_family="L2"
26264..26442
/rpt_family="(TGGG)n"
complement(28083..28216)
/rpt_family="MIR"
28255..28430
/rpt_family="MLTID"
```

```
repeat_region 28783..28913
/rpt_family="MER41D"
repeat_region 29046..29084
/rpt_family="AT-rich"
repeat_region 29111..29392
/rpt_family="AluDb"
29396..29433
/rpt_family="(CAA)n"
```

Query Match 50.8%; Score 212; DB 9; Length 182016;
Best Local Similarity 100.0%; Pred. No. 1.4e-98;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGGGAACCTGTCCGCGAGGACTCCGACTCAGGCACTGAGAGACACTGTGG 60
DB 28734 ATGGGGGAACCTGTCCGCGAGGACTCCGACTCAGGCACTGAGAGACACTGTGG 28675
QY 61 GAGATGATGAGAGACCGCCAGATCGTACCTGATCTGCGCCAGCCGCTCAC 120
DB 28674 GAGATGATGAGAGACCGCCAGATCGTACCTGATCTGCGCCAGCCGCTCAC 28615
QY 121 CCTACCTGCGCAGGCGCAAGGTGCTGTCCAGCTGAGAGAGAGAGTGTGCACAGC 180
DB 28614 CCTACCTGCGCAGGCGCAAGGTGCTGTCCAGCTGAGAGAGAGAGTGTGCACAGC 28555
QY 181 CCCGGCTCACCAACAGCCGCTATGCGGCGCG 212
DB 28554 CCCGGCTCACCAACAGCCGCTATGCGGCGCG 28523
```

RESULT 9
AC015559 187865 bp DNA linear HTG 01-MAR-2001
LOCUS
DEFINITION Homo sapiens chromosome 17 clone RP11-334C17, WORKING DRAFT
ACCESSION AC015559
VERSION AC015559.10 GI:13173726
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1 (bases 1 to 187865)
Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
Fedeerspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
Mao,J., Lam,B., Marathe,R., Miranda,M., Morehouse,A.U., Nguyen,M.,
Oefner,P., Palm,C.J., Ramirez,D., Southwick,A.M., Wilhelmy,J.,
Yu,S. and Davis,R.W.

JOURNAL
REFERENCE
2 (bases 1 to 187865)
Bruno,D., Conn,L., Dela Rosa,M., Fedeerspiel,N., Foreman,P.,
Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,
Morehouse,A.U., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,
Yu,S. and Davis,R.W.

TITLE
JOURNAL
COMMENT
On Mar 1, 2001 this sequence version replaced gi:11968315.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center

Center code: SDSTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: RP11-334C17
Center clone name: RP11-334C17
----- Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye-primer; 11k of reads
Chemistry: Dye-terminator Big Dye; 89% of reads
Assembly program: Phrap; version 0.990319

Consensus quality: 180520 bases at least Q40
 Consensus quality: 181812 bases at least Q30
 Consensus quality: 182303 bases at least Q20
 Insert size: 190560; agarose-fp
 Insert size: 186865; sum-of-contrigs
 Quality coverage: 9.0x in Q20 bases; sum-of-contrigs.
 Quality coverage: 9.2x in Q20 bases; sum-of-contrigs.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 11 contrigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contrigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 3078: contrig of 3078 bp in length
 * 3079 3178: gap of unknown length
 * 3179 6672: contrig of 3494 bp in length
 * 6673 6772: gap of unknown length
 * 6773 10097: contrig of 3325 bp in length
 * 10098 10197: gap of unknown length
 * 10197 15387: contrig of 5190 bp in length
 * 15388 15487: gap of unknown length
 * 15488 24378: contrig of 8891 bp in length
 * 24379 24478: gap of unknown length
 * 24479 34707: contrig of 10229 bp in length
 * 34708 34807: gap of unknown length
 * 34808 46756: contrig of 11949 bp in length
 * 46757 46856: gap of unknown length
 * 46857 67497: contrig of 20641 bp in length
 * 67498 67597: gap of unknown length
 * 67598 91511: contrig of 23914 bp in length
 * 91512 91611: gap of unknown length
 * 91612 118431: contrig of 26820 bp in length
 * 118432 118532: gap of unknown length
 * 118532 187865: contrig of 69334 bp in length.

FEATURES

source

1. 187865
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /clone_id="RP11-334C17"
 /clone_1bp="RP11 human BAC library 11"
 1. 3078
 /note="assembly_name:Contig31"
 misc_feature
 3179..6672
 /note="assembly_name:Contig32"
 misc_feature
 6773..10097
 /note="assembly_name:Contig33"
 misc_feature
 10198..15387
 /note="assembly_name:Contig34"
 misc_feature
 15488..24378
 /note="assembly_name:Contig35"
 misc_feature
 24479..34707
 /note="assembly_name:Contig36"
 misc_feature
 clone_end:SP6
 34808..46756
 /note="assembly_name:Contig37"
 misc_feature
 46857..67497
 /note="assembly_name:Contig38"
 misc_feature
 67598..91511
 /note="assembly_name:Contig39"
 misc_feature
 91612..118431
 /note="assembly_name:Contig40"
 misc_feature
 118532..187865
 /note="assembly_name:Contig41"

ORIGIN

Query Match 50.8%; Score 212; DB 2; Length 187865;
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;
 Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGCGAGGAGTCCGCACTACGCGCACTGACGAGGAGCACTGTGG 60
 Db 73791 ATGGGGGAACTGTGCGCGAGGAGTCCGCACTACGCGCACTGACGAGGAGCACTGTGG 73850
 QY 61 GAGTATATGAGAGACCAACCGGCAACAGATGTGATGCTGTGCTTCGCCCCACCGCTTACC 120
 Db 73851 GAGTATATGAGAGACCAACCGGCAACAGATGTGATGCTGTGCTTCGCCCCACCGCTTACC 73910
 QY 121 CCCATCTGGCGGCGAAGGCGAAGTGTGTGCGAGCTGAGACGAGAGAGAGTGTCTCAGC 180
 Db 73911 CCCATCTGGCGGCGAAGGCGAAGTGTGTGCGAGCTGAGACGAGAGAGAGTGTCTCAGC 73970
 QY 181 CCCGCGCTCACCAACAGCGCCATCGCGGCGG 212
 Db 73971 CCCGCGCTCACCAACAGCGCCATCGCGGCGG 74002

RESULT 10

AC109321

LOCUS

DEFINITION

AC109321

AC109321.15 GI:26665779

AC109321

HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 207822)

2 (bases 1 to 207822)

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

* 6200 6299: gap of 100 bp
* 6300 6969: contig of 670 bp in length
* 6970 7066: gap of 100 bp
* 7070 7779: contig of 710 bp in length
* 7880 7879: gap of 100 bp
* 8569 8568: contig of 689 bp in length
* 8669 9364: contig of 696 bp in length
* 9365 9464: gap of 100 bp
* 9465 10171: contig of 707 bp in length
* 10172 10271: gap of 100 bp
* 10272 10963: contig of 692 bp in length
* 10964 11063: gap of 100 bp
* 11064 11751: contig of 688 bp in length
* 11752 11851: gap of 100 bp
* 11852 12547: contig of 696 bp in length
* 12548 12647: gap of 100 bp
* 12648 13309: contig of 662 bp in length
* 13310 13409: gap of 100 bp
* 13410 14098: contig of 689 bp in length
* 14099 14198: gap of 100 bp
* 14199 14869: contig of 671 bp in length
* 14870 14969: gap of 100 bp
* 14970 15667: contig of 698 bp in length
* 15668 15767: gap of 100 bp
* 15768 16479: contig of 712 bp in length
* 16480 16579: gap of 100 bp
* 16580 17268: contig of 689 bp in length
* 17269 17368: gap of 100 bp
* 17369 18038: contig of 670 bp in length
* 18039 18138: gap of 100 bp
* 18139 18829: contig of 691 bp in length
* 18830 18929: gap of 100 bp
* 18930 19633: contig of 703 bp in length
* 19634 19733: gap of 100 bp
* 19734 20432: contig of 700 bp in length
* 20433 20532: gap of 100 bp
* 20533 21213: contig of 681 bp in length
* 21214 21313: gap of 100 bp
* 21314 21988: contig of 675 bp in length
* 21989 22088: gap of 100 bp
* 22089 22801: contig of 713 bp in length
* 22802 22901: gap of 100 bp
* 22902 23575: contig of 674 bp in length
* 23576 23675: gap of 100 bp
* 23676 24355: contig of 680 bp in length
* 24356 24455: gap of 100 bp
* 24456 25146: contig of 691 bp in length
* 25147 25246: gap of 100 bp
* 25247 25947: contig of 701 bp in length
* 25948 26047: gap of 100 bp
* 26048 26740: contig of 693 bp in length
* 26741 26840: gap of 100 bp
* 26841 27543: contig of 703 bp in length
* 27544 27643: gap of 100 bp
* 27644 28338: contig of 695 bp in length
* 28339 28438: gap of 100 bp
* 28439 29137: contig of 699 bp in length
* 29138 29237: gap of 100 bp
* 29238 29924: contig of 687 bp in length
* 29925 30024: gap of 100 bp
* 30025 30712: contig of 688 bp in length
* 30713 30812: gap of 100 bp
* 30813 31504: contig of 692 bp in length
* 31505 31604: gap of 100 bp
* 31605 32322: contig of 718 bp in length
* 32323 32422: gap of 100 bp
* 32423 33104: contig of 682 bp in length
* 33105 33204: gap of 100 bp
* 33205 33899: contig of 695 bp in length
* 33900 34690: contig of 691 bp in length
* 34691 34790: gap of 100 bp

* 34791 35481: contig of 691 bp in length
* 35482 35581: gap of 100 bp
* 35582 36277: contig of 696 bp in length
* 36278 36377: gap of 100 bp
* 36378 37072: contig of 695 bp in length
* 37073 37172: gap of 100 bp
* 37173 37856: contig of 684 bp in length
* 37857 37956: gap of 100 bp
* 37957 38652: contig of 696 bp in length
* 38653 38752: gap of 100 bp
* 38753 39440: contig of 688 bp in length
* 39441 39541: gap of 100 bp
* 39541 40218: contig of 678 bp in length
* 40219 40318: gap of 100 bp
* 40319 41015: contig of 697 bp in length
* 41016 41115: gap of 100 bp
* 41116 41820: contig of 705 bp in length
* 41821 41920: gap of 100 bp
* 41921 42600: contig of 680 bp in length
* 42601 42700: gap of 100 bp
* 42701 43386: contig of 686 bp in length
* 43387 43486: gap of 100 bp
* 43487 44179: contig of 693 bp in length
* 44180 44279: gap of 100 bp
* 44280 44939: contig of 660 bp in length
* 44940 45039: gap of 100 bp
* 45040 45703: contig of 664 bp in length
* 45704 45803: gap of 100 bp
* 45804 46470: contig of 667 bp in length
* 46471 46570: gap of 100 bp
* 46571 47264: contig of 694 bp in length
* 47265 47364: gap of 100 bp
* 47365 48069: contig of 705 bp in length
* 48070 48169: gap of 100 bp
* 48170 48847: contig of 678 bp in length
* 48848 48947: gap of 100 bp
* 48948 49636: contig of 669 bp in length
* 49637 49736: gap of 100 bp
* 49737 50437: contig of 701 bp in length
* 50438 50537: gap of 100 bp
* 50538 51213: contig of 676 bp in length
* 51214 51313: gap of 100 bp
* 51314 51999: contig of 666 bp in length
* 52000 52099: gap of 100 bp
* 52099 52792: contig of 693 bp in length
* 52793 52892: gap of 100 bp
* 52893 53594: contig of 702 bp in length
* 53595 53694: gap of 100 bp
* 53695 54377: contig of 683 bp in length
* 54378 54477: gap of 100 bp
* 54478 55173: contig of 696 bp in length.

FEATURES
source location/Qualifiers
1. .55173

Query Match 49.2%; Score 205; DB 2; Length 55173;
Best Local Similarity 100.0%; Pred. No. 6; 9e-95;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 GGGCACTTGCTGGATTGCTGAAGCTGAGGGAAGAACGGGGCCATGCTCCTCTGAG 270
Db 20167 GGGCACTTGCTGGATTGCTGAAGCTGAGGGAAGAACGGGGCCATGCTCCTCTGAG 20108

QY 271 AGCGTAATTCCAAACCTGACCTGACACCCGTGACACCGGCGTGCAGCTGATGTT 330
Db 20107 AGCGTAATTCCAAACCTGACCTGACACCCGTGACACCGGCGTGCAGCTGATGTT 20048

QY 331 GACTTCAGTACTTTCAGGCTGAGAGCTCCGACTTTGACGTTTGGACGACCTTCTAGG 390
Db 20047 GACTTCAGTACTTTCAGGCTGAGAGCTCCGACTTTGACGTTTGGACGACCTTCTAGG 19988

QY 391 AACCTGAGGCTCTGTGTAACCCGAG 415
Db 19987 AACCTGAGGCTCTGTGTAACCCGAG 19963

RESULT 12
AC137736
LOCUS
DEFINITION
AC137736 Homo sapiens chromosome 17 clone RP13-365D24 map 17, LOW-PASS
SEQUENCE SAMPLING.
AC137736
AC137736.1 GI:25815500
HMG; HTGS_PHASE0
SOURCE
Homo sapiens (human)
ORGANISM
REFERENCE
Birken, B., Nussbaum, C. and Lander, E.
1 (bases 1 to 74301)
Homo sapiens chromosome 17, clone RP13-365D24
2 (bases 1 to 74301)
Unpublished
REFERENCE
Birken, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhassira, B.,
Camataia, J., Chang, J., Chazaro, B., Choquet, Y., Collymore, A.,
Cook, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagoe, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,
Mathews, C., McCarthy, M., Meldrum, J., Meneses, L., Mhova, T.,
Ming, V., Murphy, T., Nguyen, C., Nicol, R., North, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Rette, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Totham, K.,
Travers, M., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (28-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L28728
Center clone name: 365_D_24
*
* NOTE: This record contains 91 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 707: contig of 707 bp in length
* 708 807: gap of 100 bp
* 808 1540: contig of 733 bp in length
* 1541 1640: gap of 100 bp
* 1641 2344: contig of 704 bp in length
* 2345 2444: gap of 100 bp
* 2445 3156: contig of 712 bp in length
* 3157 3256: gap of 100 bp
* 3257 3978: contig of 722 bp in length
* 3979 4078: gap of 100 bp

* 4079 4781: contig of 703 bp in length
* 4782 4881: gap of 100 bp
* 4882 5594: contig of 713 bp in length
* 5595 5694: gap of 100 bp
* 5695 6405: contig of 711 bp in length
* 6406 6505: gap of 100 bp
* 6506 7197: contig of 692 bp in length
* 7198 7297: gap of 100 bp
* 7298 8015: contig of 718 bp in length
* 8016 8115: gap of 100 bp
* 8116 8818: contig of 703 bp in length
* 8819 8918: gap of 100 bp
* 8919 9653: contig of 735 bp in length
* 9654 9753: gap of 100 bp
* 9754 10454: contig of 701 bp in length
* 10455 10554: gap of 100 bp
* 10555 11277: contig of 723 bp in length
* 11278 11377: gap of 100 bp
* 11378 12081: contig of 704 bp in length
* 12082 12181: gap of 100 bp
* 12182 12913: contig of 732 bp in length
* 12914 13013: gap of 100 bp
* 13014 13742: contig of 725 bp in length
* 13743 13842: gap of 100 bp
* 13843 14560: contig of 718 bp in length
* 14561 14660: gap of 100 bp
* 14661 15382: contig of 722 bp in length
* 15383 15482: gap of 100 bp
* 15483 16198: contig of 716 bp in length
* 16199 16298: gap of 100 bp
* 16299 17006: contig of 708 bp in length
* 17007 17106: gap of 100 bp
* 17107 17817: contig of 711 bp in length
* 17818 17917: gap of 100 bp
* 17918 18642: contig of 725 bp in length
* 18643 18742: gap of 100 bp
* 18743 19458: contig of 716 bp in length
* 19459 19558: gap of 100 bp
* 19559 20264: contig of 706 bp in length
* 20265 20364: gap of 100 bp
* 20365 21103: contig of 739 bp in length
* 21104 21203: gap of 100 bp
* 21204 21910: contig of 707 bp in length
* 21911 22010: gap of 100 bp
* 22011 22731: contig of 721 bp in length
* 22732 22831: gap of 100 bp
* 22832 23556: contig of 725 bp in length
* 23557 23656: gap of 100 bp
* 23657 24373: contig of 717 bp in length
* 24374 24473: gap of 100 bp
* 24474 25166: contig of 693 bp in length
* 25167 25266: gap of 100 bp
* 25267 25949: contig of 683 bp in length
* 25950 26049: gap of 100 bp
* 26050 26763: contig of 714 bp in length
* 26764 26863: gap of 100 bp
* 26864 27583: contig of 720 bp in length
* 27584 27683: gap of 100 bp
* 27684 28421: contig of 738 bp in length
* 28422 28521: gap of 100 bp
* 28522 29223: contig of 702 bp in length
* 29224 29323: gap of 100 bp
* 29324 30031: contig of 708 bp in length
* 30032 30131: gap of 100 bp
* 30132 30859: contig of 728 bp in length
* 30860 30959: gap of 100 bp
* 30960 31693: contig of 734 bp in length
* 31694 31793: gap of 100 bp
* 31794 32516: contig of 723 bp in length
* 32517 32616: gap of 100 bp
* 32617 33338: contig of 722 bp in length
* 33339 33438: gap of 100 bp
* 33439 34162: contig of 724 bp in length

```
*
* 34163 34262: gap of 100 bp
* 34263 34982: contig of 720 bp in length
* 34983 35082: gap of 100 bp
* 35083 35785: contig of 703 bp in length
* 35786 35885: gap of 100 bp
* 35886 36612: contig of 727 bp in length
* 36613 36712: gap of 100 bp
* 36713 37460: contig of 748 bp in length
* 37461 37560: gap of 100 bp
* 37561 38293: contig of 733 bp in length
* 38294 38393: gap of 100 bp
* 38394 39114: contig of 721 bp in length
* 39115 39214: gap of 100 bp
* 39215 39943: contig of 729 bp in length
* 39944 40043: gap of 100 bp
* 40044 40763: contig of 720 bp in length
* 40764 40863: gap of 100 bp
* 40864 41562: contig of 699 bp in length
* 41563 41662: gap of 100 bp
* 41663 42386: contig of 724 bp in length
* 42387 42486: gap of 100 bp
* 42487 43202: contig of 716 bp in length
* 43203 43302: gap of 100 bp
* 43303 44007: contig of 705 bp in length
* 44008 44107: gap of 100 bp
* 44108 44794: contig of 687 bp in length
* 44795 44894: gap of 100 bp
* 44895 45581: contig of 687 bp in length
* 45582 45681: gap of 100 bp
* 45682 46463: contig of 682 bp in length
* 46464 47166: gap of 100 bp
* 47167 47266: gap of 100 bp
* 47267 48002: contig of 736 bp in length
* 48003 48102: gap of 100 bp
* 48103 48831: contig of 729 bp in length
* 48832 48931: gap of 100 bp
* 48932 49651: contig of 720 bp in length
* 49652 49751: gap of 100 bp
* 49752 50488: contig of 737 bp in length
* 50489 50588: gap of 100 bp
* 50589 51295: contig of 707 bp in length
* 51296 51395: gap of 100 bp
* 51396 52123: contig of 728 bp in length
* 52124 52223: gap of 100 bp
* 52224 52935: contig of 712 bp in length
* 52936 53035: gap of 100 bp
* 53036 53743: contig of 708 bp in length
* 53744 53843: gap of 100 bp
* 53844 54538: contig of 695 bp in length
* 54539 55354: contig of 716 bp in length
* 55355 56159: gap of 100 bp
* 56159 56259: gap of 100 bp
* 56260 56986: contig of 727 bp in length
* 56987 57807: gap of 100 bp
* 57808 57807: contig of 721 bp in length

Query Match 38.6%; Score 161; DB 2; Length 74301;
Best Local Similarity 100.0%; Pred. No. 3.9e-72;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 24305 CCTACTGCGCCAGGCCAAGTGTCTGCGCAGTGAACA 24345

RESULT 13
AF363457
LOCUS 3772 bp mRNA linear ROD 14-AUG-2001
DEFINITION Mus musculus BIMP2 (Bimp2) mRNA, complete cds.
ACCESSION AF363457
VERSION AF363457.1 GI:14326096
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3772)
McAllister-Lucas,L.M., Inohara,N., Lucas,P.C., Ruland,U.,
Benito,A., Li,Q., Chen,S., Chen,F.F., Yamaoka,S., Verma,I.M.,
Mak,T.W. and Nunez,G.
Bimp1, a MAPK family member linking protein kinase C activation to
Bcl10-mediated NF-kappaB induction
J. Biol. Chem. 276 (33), 30589-30597 (2001)
21391892
JOURNAL
MEDLINE 11387339
PUBMED 2 (bases 1 to 3772)
REFERENCE McAllister-Lucas,L.M., Lucas,P.C., Inohara,N., Chen,S., Chen,F.F.
and Nunez,G.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2001) Department of Pathology, University of
Michigan Medical School, 1500 E. Medical Center Dr., Ann Arbor, MI
48109, USA

FEATURES
SOURCE
Location/Qualifiers
1..3772
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..3772
/gene="Bimp2"
304..3303
/gene="Bimp2"
/codon_start=1
/product="BIMP2"
/db_xref="gi:14326097"
/db_xref="gi:14360137.1"
/translation="MAELCRKDSLTALDEMDMDNLSHRCRIVQSCSRITPYLR
QAKVLQSDDEEILHSRFTNSARVGHLDLPAKGNKAIATLESUKFNPDPVYL
VTGLOSIDIDFISGLMETSRLTECLAGLSLKEAEKQKAVLELRQQLKEL
GLAEHAEGLRQLEVDHSMRREYVTHFEVTLKIDEMLSLHYSNLRKEKLAATR
CHSIOELVLYKQELORASIVSCERESRESLMAASNLBQGEINLRKEKELRS
MPSIVVEDILQSDIDARERESKQSLNDRIRHSLERAVAAERQCVWEKEQTLQPR
KTQVCELYKEKMTLQGVAELOKRPDQATYARDPROMETISQRLVERDALRRVYEL
TEQVCELTQRLQAEAPGPKQAGARQELCLGKQRLVMAHVCPDDSDCSLSS
TESRLMDLNTSSREOMDSFRSSSPMPQOQSLYKRVADFLDEPBLSPREYLEMR
LOGATVMDTDLDLEFEMTDGADLSOTEDSLQGSRLNVSSEVPRRPAKTLISQV
TVLANQEPALLAQGVIGVIGNTLGIPIRHVTEGSADEALAPGOIMVVDYKRPESL
RATENNTLEQVGLRRVNSCTLSVKINTEGYNTLLODDAKYVTSGBDFYIRVNL
AMORGEDELQTCNDILHVDTFMQGSCWHAHVNTYKIDMEPTIPYISOQOQ
LLALLODMTORCTVPRKPGPKQLVLRVSVKXAVSLTSSPQSQMDSEKEBGS
VCFNESCEFLAPVTLVPHRPARPRPYLPVRLVGRILGKLCILQFKQCSAEYLS
OEVAWTSORGDITQEGESIGDHMITTHAVESIMNSHTALLDVRSLVLRNDM
PFIITHVYNEKTKAKLKGKLRIGSSSEFQLEVARQEBGLDVRPCLYSLAPDSWS
DUDLSLSCVRLALDEQKRVWTSPPC"
```

```
ORIGIN
Query Match 7.9%; Score 33; DB 10; Length 3772;
Best Local Similarity 100.0%; Pred. No. 9.1e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 115 CTCACCCCTACTGCGCCAGGCCAAGTGTCTG 147
418 CTCACCCCTACTGCGCCAGGCCAAGTGTCTG 450
```

RESULT 14

BC029102 3995 bp mRNA linear ROD 07-OCT-2003
 LOCUS BC029102
 DEFINITION Mus musculus caspase recruitment domain family, member 14, mRNA
 (CDNA clone MGC:28122 IMAGE:3979883), complete cds.
 ACCESSION BC029102
 VERSION BC029102.1 GI:22137687
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3995)
 Strausberg, R.L., Feingold, F.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., Stopleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Umedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfield, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22386257
 12477932
 2 (bases 1 to 3995)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 REMARK COMMENT
 CONTACT: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at: <http://image.llnl.gov>
 Series: IRAK Plate: 36 Row: e Column: 1
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 18700027.
 Location/Qualifiers
 1. 3995
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:28122 IMAGE:3979883"
 /tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
 /clone_lib="NCI_CGAP_Mam1"
 /lab_host="DH10B"

gene
 1. 3995
 /gene="Card14"
 /note="synonyms: Bimp2, CARMA2"
 /db_xref="LocusID:170720"
 /db_xref="MGI:2386258"
 325..3324
 /codon_start=1
 /product="caspase recruitment domain family, member 14"
 /protein_id="AAH29102.1"
 /db_xref="GI:22137688"
 /db_xref="LocusID:170720"
 /translation="MAELRMDSTPLALDEEMLMWLESHRCRIVOSICPSRLPYIR
 OKAVUQDLDEEBILHSRRTNANRGHILLDLKAGKGAIAFESLTFHPDYVTL
 VTELQDIDPSTYSGLMERSKLTJECLAGAISTQELADQKQATLRLRCCQKRL
 GLAEARAEGRQLEVDHSMKEVSHFEHVEIKQEMNLISHTSNALREKELKLR
 CHLORELYLVKQELQRAVLVSCESESERSTKMSNLNPEQEEELNRLKEENKLS
 MPESLVEKILDSLBARESKEOELVDRHSIRERAVARROOKOYKEKECTLLQFR
 KYQVDELYKEKNTMIGOVALLQKRDQAYTARDAQEISQRIYVEXALRRVEL
 TEQVCELRQLQRLQLEAPGKQEGARELCIRGKQRLVRMAVCPDPDSCSLIS
 TESRLWDLNLSREQMSFRSSPEPQSLYKRVADLEDESLSPFVLEMR
 LQATVDPTDIDLEFEMIDGADLSQTEDELSSESLNBSVPRRPARIKLSQV
 TVLAFQDALLBQIGVINGLGIPIHRTVPSADENALRGTOIMADYPTKSL
 RALERTTLEQANGLRRVNSCVLSKINTGYKXLLQDLAKVTSQDSFYIVNL
 AMQKGDGELQTHCNDILHVTDMFGKSCAHANRYPTKMDMERGTLPTNSQAOQ
 LIALIDMQRCTVPRPGQKLVIVSVKAAVPLTSSFDQSQMSGKEEGPS
 VCPWSSCFGLAPYTLVHPHPPARPPVFLVGRILGKICLQGRKQCSAEYLS
 OEVYAWSSQGDIOGESIGDHHTVTRAVSLNMSHTALIDVRLSVYLRMDM
 PRITIVSNENKAKLRKGLRHSSESQFLVARQEGELDRVPCVLSLAPDSMS
 DDLSTLCRLALADQKQVWTEPSC"
 1201..1617
 /note="ERM; Region: Ezrin/radixin/moesin family. This family of proteins contain a band 4.1 domain (pfam00373), at their amino terminus. This family represents the rest of these proteins"
 /db_xref="CD:pfam00769"
 2059..2280
 /note="PDZ; Region: Domain present in PSD-95, Dlg, and ZO-1/2. Also called DHR (Dlg homologous region) or GGF (relatively well conserved tetrapeptide in these domains). Some PDZs have been shown to bind C-terminal polypeptides"
 /db_xref="CD:smat00228"
 ORIGIN
 Query Match 7.9%; Score 33; DB 10; Length 3995;
 Best Local Similarity 100.0%; Pred. No. 9,1e-06;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 115 CTCACCCCTACTCTGCGCCAGGCAAGTGCTG 147
 Db 439 CTCACCCCTACTCTGCGCCAGGCAAGTGCTG 471
 RESULT 15
 LOCUS AL645911
 DEFINITION Mouse DNA sequence from clone RP23-25M3 on chromosome 11, complete sequence.
 ACCESSION AL645911
 VERSION AL645911.14 GI:22204307
 KEYWORDS HMG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 244653)
 Tiomans, A.
 Direct Submission
 Submitted (03-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquer@sanger.ac.uk
 On Aug 11, 2002 this sequence version replaced gi:21531055.
 COMMENT
 ----- Genome Center

Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-25M3 is from the RP23-25 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBac3.6.

FEATURES

SOURCE

1..244653
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-25M3"
 /clone_id="RP23-23"

ORIGIN

Query Match 7.9%; Score 33; DB 10; Length 244653;

Best Local Similarity 100.0%; Pred. No. 5.6e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 115 CTCACCCCTACTGCGCCAGCCAGGATGCTG 147
 |||||
 Db 24155 CTCACCCCTACTGCGCCAGCCAGGATGCTG 24187

Search completed: March 22, 2004, 01:25:26
 Job time : 3004 secs

XX Pawlowski K, Reed JC, Godzik A;
 XX MPI; 2003-288137/28.
 DR P-PSDB; ABG76062.

XX New isolated CARD-containing nucleic acids, useful for the diagnosis and
 PT treatment of disorders with aberrant expression or activity of the CARD-
 PT containing polypeptide, such as cancer, stroke, arthritis, heart failure
 PT and AIDS.

XX Claim 1; Fig 3; 34pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding a
 CC caspase recruitment domain (CARD) containing polypeptide. CARD containing
 CC polypeptides are involved in apoptosis (as caspase activators and caspase
 CC inhibitors), cell adhesion, inflammation and cytokine receptor
 CC signalling. The methods and compositions of the present invention are
 CC useful for the diagnosis and treatment of disorders associated with the
 CC aberrant expression or activity of the CARD containing polypeptide such
 CC as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,
 CC hamartoma, leukaemia, lymphoma, keratinocyte hyperplasia, neoplasia,
 CC benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis,
 CC restenosis, allergies, inflammatory diseases such as arthritis, lupus,
 CC Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus
 CC host disease, sepsis, abnormal cell death diseases such as stroke,
 CC myocardial infarction, heart failure, neurodegenerative diseases like
 CC Parkinson's disease and Alzheimer's disease, and HIV infection. The
 CC present sequence represents cDNA encoding the human caspase recruitment
 CC domain containing protein, CARD-12X

XX Sequence 417 BP; 84 A; 131 C; 128 G; 72 T; 0 U; 2 Other;

XX Query Match 100.0%; Score 417; DB 7; Length 417;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-197;

XX Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 ATGGGGGAACCTGTGCGCGAGGACTCCGCACTACAGGCACTGAGAGAGACACTGTGG 60
 DB 1 ATGGGGGAACCTGTGCGCGAGGACTCCGCACTACAGGCACTGAGAGAGACACTGTGG 60
 QY 61 GAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 DB 61 GAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 121 CCTACTCTGCGCCAGGCCAGAGTGTGTGCCAGCTGAGAGAGAGAGAGAGAGAGAGAG 180
 DB 121 CCTACTCTGCGCCAGGCCAGAGTGTGTGCCAGCTGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 181 CCGCGGCTACCAACAGCGCCATGCGGGCGGGGCACTTGCTGAGATTGCTGAAGACTCGA 240
 DB 181 CCGCGGCTACCAACAGCGCCATGCGGGCGGGGCACTTGCTGAGATTGCTGAAGACTCGA 240
 QY 241 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 DB 241 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 301 ACCCTGCTCACCGGGCTGAGCTGATGTGACTTCACTTAAGAGAGAGAGAGAGAGAGAG 360
 DB 301 ACCCTGCTCACCGGGCTGAGCTGATGTGACTTCACTTAAGAGAGAGAGAGAGAGAGAG 360
 QY 361 GACTTTGACGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417
 DB 361 GACTTTGACGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417

RESULT 2
 AAD13448
 ID AAD13448 standard; DNA; 3417 BP.
 AC AAD13448;
 XX
 DT 06-NOV-2001 (first entry)

XX Human caspase recruitment domain-14 (CARD-14) genomic DNA.
 DE
 XX Human; caspase recruitment domain-14; CARD-14; chromosome 17;
 XX nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;
 KW autoimmune disorder; systemic lupus erythematosus; neurological disorder;
 KW Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;
 KW haematological disorder; myelodysplastic syndrome; myocardial infarction;
 KW stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;
 KW cell signalling disorder; cytostatic; immunosuppressive; nootropic;
 KW neuroprotective; antiviral; antibacterial; ds.

XX Homo sapiens.

Key	Location/Qualifiers
CDS	1..3417
FT	/tag= a
FT	/product= "Human caspase recruitment domain-14"
FT	/transl_except= (pos:2098..2100, aa:Cys)
FT	/note= "CDS does not include stop codon"
FT	/partial
FT	1..211
FT	/tag= b
FT	/number= 1
FT	212..415
FT	/tag= c
FT	/number= 2
FT	416..741
FT	/tag= d
FT	/number= 3
FT	742..909
FT	/tag= e
FT	/number= 4
FT	910..1029
FT	/tag= f
FT	/number= 5
FT	1030..1209
FT	/tag= g
FT	/number= 6
FT	1210..1359
FT	/tag= h
FT	/number= 7
FT	1360..1476
FT	/tag= i
FT	/number= 8
FT	1477..1619
FT	/tag= j
FT	/number= 9
FT	1620..1714
FT	/tag= k
FT	/number= 10
FT	1715..1778
FT	/tag= l
FT	/number= 11
FT	1779..2001
FT	/tag= m
FT	/number= 12
FT	2002..2128
FT	/tag= n
FT	/number= 13
FT	2129..2369
FT	/tag= o
FT	/number= 14
FT	2370..2433
FT	/tag= p
FT	/number= 15
FT	2434..2548
FT	/tag= q
FT	/number= 16
FT	2549..2719
FT	/tag= r
FT	/number= 17
FT	2720..2841

FT exon

```

FT      /*tag= s
FT      /number= 18
FT      exon      2842. .2957
FT      /*tag= t
FT      /number= 19
FT      exon      2958. .3138
FT      /*tag= u
FT      /number= 20
FT      exon      3139. .3417
FT      /*tag= v
FT      /number= 21
XX      WO200159065-A2.
XX      16-AUG-2001.
XX      22-JAN-2001; 2001WO-US002087.
XX      09-FEB-2000; 2000US-0181159P.
XX      (MILL-) MILLENNIUM PHARM INC.
XX      Bertin J;
XX      WPI; 2001-497073/54.
XX      P-PSDB; AAE07165.
XX      An isolated caspase recruitment domain polypeptide useful for regulating
XX      growth and cell death and useful for the treatment of cancer.
XX      Disclosure; Fig 2A-2C; 109pp; English.
XX      The present sequence is human caspase recruitment domain-14 (CARD-14)
XX      DNA. The CARD-14 gene is located on chromosome 17. The CARD-14 is used
XX      for the detection of modulators that modulates the ability of CARD-14 to
XX      bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation of
XX      nuclear factor-kappa B (NF-kB). The CARD-14 is useful for regulating
XX      growth and cell death and useful for the treatment of cancer. It is also
XX      useful for the treatment of autoimmune disorders (e.g., systemic lupus
XX      erythematosus), neurological disorders (e.g., Alzheimer's and Parkinson's
XX      disease), inflammatory disorders, haematological disorders (e.g., anaemia,
XX      myelodysplastic syndromes), myocardial infarctions, strokes, immune
XX      disorders (e.g., Crohn's disease, allergic rhinitis), cell signalling
XX      disorders and certain viral and bacterial infections
XX      Sequence 3417 BP; 739 A; 1039 C; 1102 G; 537 T; 0 U; 0 Other;
SQ
Query Match      99.5%; Score 415; DB 4; Length 3417;
Best Local Similarity 100.0%; Pred. NO. 1e-196;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      301 ACCCTGTACCGGGGCTGCAGGCGTATGTTGACTTCACTTAGACGGGTGAGAGCTCC 360
Qy      361 GACTTTGACGGGTTTGGCAGGCACTTTAGAAACCTGAGGCTCTGTGTAACCCAG 415
Db      361 GACTTTGACGGGTTTGGCAGGCACTTTAGAAACCTGAGGCTCTGTGTAACCCAG 415
XX      RESULT 3
XX      ADB62161
XX      ID ADB62161 standard; cDNA; 3766 BP.
XX      AC ADB62161;
XX      DT 04-DEC-2003 (first entry)
XX      DE Human cDNA encoding clone CTONG20007660.
XX      KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
XX      KW tissue regeneration; cell regeneration; membrane protein;
XX      KW signal transduction-related protein; transcription-related protein;
XX      KW osteoporosis; neurological disease; cancer; tumour.
XX      OS Homo sapiens.
XX      FH Key Location/Qualifiers
XX      FT CDS 192..1448
XX      FT /*tag= a
XX      FT /product= "Clone CTONG20007660 protein"
XX      EN EP1308459-A2.
XX      PD 07-MAY-2003.
XX      PF 28-MAR-2002; 2002EP-00007401.
XX      PR 05-NOV-2001; 2001JP-00379298.
XX      PR 25-JAN-2002; 2002US-00350978.
XX      PA (HELI-) HELIX RES INST.
XX      PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX      PI Iosogi T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX      PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX      PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX      DR WPI; 2003-450961/43.
XX      DR P-PSDB; ADB64131.
XX      PT New polynucleotides and polypeptides, useful for developing a diagnostic
XX      PT marker or medicines for regulation of their expression and activity, or
XX      PT as targets of gene therapy.
XX      PS Claim 1; Page; 222pp; English.
XX      The invention discloses a polynucleotide comprising a sequence selected
XX      CC from 1970 fully defined nucleotide sequences which encode novel
XX      CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX      CC or its partial peptide, an antibody binding to the polypeptide or peptide
XX      CC of the polynucleotide, immunologically assaying the polypeptide or
XX      CC peptide of the polynucleotide by contacting the polypeptide or peptide
XX      CC with the antibody of the encoded protein, and observing the binding
XX      CC between the two, a transformant carrying the polynucleotide in an
XX      CC expressible manner and an antisense polynucleotide. The oligonucleotide
XX      CC is useful as a primer for synthesizing the polynucleotide, or as a probe
XX      CC for detecting the polynucleotide. The polynucleotides and encoded
XX      CC proteins are useful as pharmaceutical agents and many disease-related
XX      CC genes may be included in them, for developing a diagnostic marker or
XX      CC medicines for regulation of their expression and activity, or as targets
XX      CC of gene therapy. The genes are involved in tissue and/or cell
XX      CC regeneration. Membrane proteins, signal transduction-related proteins,
XX      CC transcription-related proteins, disease-related proteins and genes
XX      CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
XX      CC neurological diseases, cancer, tumours. The cDNA may be used to regulate

```


CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.

XX Sequence 3766 BP, 764 A, 1140 C, 1217 G, 645 T, 0 U, 0 Other;

Query Match 84.2%; Score 351; DB 9; Length 3766;

Best Local Similarity 100.0%; Pred. No. 8.2e-165;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGGAACTGTGCGCGAGGACTCCGCACTACGCGCACTGAGCAGAGACACTGTGG 60
DB 192 ATGGGGAACTGTGCGCGAGGACTCCGCACTACGCGCACTGAGCAGAGACACTGTGG 251
QY 61 GAGATGATGAGAGCCACCGCCACAGGATGTAGGCTGATGCCCGCCAGCCCTCACC 120
DB 252 GAGATGATGAGAGCCACCGCCACAGGATGTAGGCTGATGCCCGCCAGCCCTCACC 311
QY 121 CCTACTGCGCGCGCAAGTGTGTGTGCGCACTGAGCAGAGAGGTGTGCACAGC 180
DB 312 CCTACTGCGCGCGCAAGTGTGTGTGCGCACTGAGCAGAGAGGTGTGCACAGC 371
QY 181 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTGTGTGATTGTCTGAAAGCTCGA 240
DB 372 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTGTGTGATTGTCTGAAAGCTCGA 431
QY 241 GGGAGAAACGGGGCGCATGCTTCTGTGAGAGCTGTGAAAGTTCACAACTGACGCTTAC 300
DB 432 GGGAGAAACGGGGCGCATGCTTCTGTGAGAGCTGTGAAAGTTCACAACTGACGCTTAC 491
QY 301 ACCCTGTGACCGGGCTGACGCGTGAATGTGACTTCACTTAAGCGGT 351
DB 492 ACCCTGTGACCGGGCTGACGCGTGAATGTGACTTCACTTAAGCGGT 542
```

RESULT 4

AD13447 standard, cDNA, 3931 BP.

AD13447;

06-NOV-2001 (first entry)

Human caspase recruitment domain-14 (CARD-14) cDNA.

Human; caspase recruitment domain-14; CARD-14; chromosome 17;
nuclear factor-kappa B; NF-kB; cell growth; cell death; cancer; therapy;
autoimmune disorder; systemic lupus erythematosus; neurological disorder;
Alzheimer's disease; Parkinson's disease; inflammatory disorder; anaemia;
haematological disorder; myelodysplastic syndrome; myocardial infarction;
stroke; immune disorder; Crohn's disease; allergic rhinitis; infection;
cell signalling disorder; cytostatic; immunosuppressive; nootropic;
neuroprotective; antiviral; antibacterial; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 207..3221

FT /product= "Human caspase recruitment domain-14"
FT /note= "Residues 207-3218 is specifically claimed as SEQ
ID NO 3 in claim 9 of the specification"

W0200159065-A2.

16-AUG-2001.

22-JAN-2001; 2001WO-US002087.

09-FEB-2000; 2000US-0181159P.

(MILL-) MILLENNIUM PHARM INC.

XX Bertin U;
PI WPI; 2001-497073/54.
DR P-PsDB; AAB07164.

PT An isolated caspase recruitment domain polypeptide useful for regulating
PT growth and cell death and useful for the treatment of cancer.
PS Claim 9; Fig 1A-1E; 109pp; English.

CC The present sequence is human caspase recruitment domain-14 (CARD-14)
CC cDNA. The CARD-14 gene is located on chromosome 17. The CARD-14 is used
CC for the detection of modulators that modulates the ability of CARD-14 to
CC bind to Bcl-10 and stimulates phosphorylation of Bcl-10 or activation of
CC nuclear factor-kappa B (NF-kB). The CARD-14 is useful for regulating
CC growth and cell death and useful for the treatment of cancer. It is also
CC useful for the treatment of autoimmune disorders (e.g., systemic lupus
CC erythematosus), neurological disorders (e.g., Alzheimer's and Parkinson's
CC disease), inflammatory disorders, haematological disorders (e.g., anaemia,
CC myelodysplastic syndromes), myocardial infarctions, strokes, immune
CC disorders (e.g., Crohn's disease, allergic rhinitis), cell signalling
CC disorders and certain viral and bacterial infections

XX Sequence 3931 BP, 799 A, 1187 C, 1263 G, 682 T, 0 U, 0 Other;

Query Match 84.2%; Score 351; DB 4; Length 3931;

Best Local Similarity 100.0%; Pred. No. 8.1e-165;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGGAACTGTGCGCGAGGACTCCGCACTACGCGCACTGAGCAGAGACACTGTGG 60
DB 207 ATGGGGAACTGTGCGCGAGGACTCCGCACTACGCGCACTGAGCAGAGACACTGTGG 266
QY 61 GAGATGATGAGAGCCACCGCCACAGGATGTAGGCTGATGCCCGCCAGCCCTCACC 120
DB 267 GAGATGATGAGAGCCACCGCCACAGGATGTAGGCTGATGCCCGCCAGCCCTCACC 326
QY 121 CCTACTGCGCGCGCAAGTGTGTGTGCGCACTGAGCAGAGAGGTGTGCACAGC 180
DB 327 CCTACTGCGCGCGCAAGTGTGTGTGCGCACTGAGCAGAGAGGTGTGCACAGC 386
QY 181 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTGTGTGATTGTCTGAAAGCTCGA 240
DB 387 CCGCGGCTCACCAACAGCGCATGCGGGCGGCGCACTGTGTGATTGTCTGAAAGCTCGA 446
QY 241 GGGAGAAACGGGGCGCATGCTTCTGTGAGAGCTGTGAAAGTTCACAACTGACGCTTAC 300
DB 447 GGGAGAAACGGGGCGCATGCTTCTGTGAGAGCTGTGAAAGTTCACAACTGACGCTTAC 506
QY 301 ACCCTGTGACCGGGCTGACGCGTGAATGTGACTTCACTTAAGCGGT 351
DB 507 ACCCTGTGACCGGGCTGACGCGTGAATGTGACTTCACTTAAGCGGT 557
```

RESULT 5

ABX11449 standard, cDNA, 1141 BP.

ABX11449;

09-MAY-2003 (first entry)

Human CARD-12X expressed sequence tag, GI10316320.

Human; ss; CARD; EST; caspase recruitment domain; apoptosis; lupus;
cell adhesion; inflammation; cytokine receptor signalling; glioma;
carcinoma; adenocarcinoma; hamartoma; leukaemia; lymphoma; melanoma;
CARD-containing polypeptide associated disorder; sarcoma; neoplasia;
keratinocyte hyperplasia; keloid; benign prostatic hypertrophy; sepsis;
inflammatory hyperplasia; fibrosis; restenosis; allergy; arthritis;
Sjogren's syndrome; Crohn's disease; ulcerative colitis; stroke; cancer;
graft versus host disease; abnormal cell death disease; heart failure;

KW myocardial infarction; neurodegenerative disease; Parkinson's disease;
XX Alzheimer's disease; HIV; CARD-12X; caspase activator; caspase inhibitor;
XX expressed sequence tag.
XX Homo sapiens.
XX OS
XX PN US2002164703-A1.
XX PD
XX PF 07-NOV-2002.
XX PP 19-DEC-2001; 2001US-00032159.
XX PR 21-DEC-2000; 2000US-0257457P.
XX PA (PAML/) PAWLOWSKI K.
XX PA (REED/) REED J C.
XX PA (GODZ/) GODZIK A.
XX PI Pawlowski K, Reed JC, Godzik A;
XX DR WPI; 2003-288137/28.
XX PT New isolated CARD-containing nucleic acids, useful for the diagnosis and
XX treatment of disorders with aberrant expression or activity of the CARD-
XX containing polypeptide, such as cancer, stroke, arthritis, heart failure
XX and AIDS.
XX PS Disclosure: Fig 7; 34pp; English.
XX CC The invention relates to an isolated nucleic acid molecule encoding a
XX caspase recruitment domain (CARD) containing polypeptide. CARD containing
XX polypeptides are involved in apoptosis (as caspase activators and caspase
XX inhibitors), cell adhesion, inflammation and cytokine receptor
XX signalling. The methods and compositions of the present invention are
XX useful for the diagnosis and treatment of disorders associated with the
XX aberrant expression or activity of the CARD containing polypeptide such
XX as cancer, glioma, carcinoma, adenocarcinoma, sarcoma, melanoma, keloid,
XX hamartoma, leukemia, lymphoma, keratinocyte hyperplasia, neoplasia,
XX benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis,
XX resectosis, allergies, inflammatory diseases such as arthritis, lupus,
XX Sjogren's syndrome, Crohn's disease, ulcerative colitis, graft versus
XX host disease, sepsis, abnormal cell death diseases such as stroke,
XX myocardial infarction, heart failure, neurodegenerative diseases like
XX Parkinson's disease and Alzheimer's disease, and HIV infection. The
XX present sequence represents the human caspase recruitment domain
XX containing protein, CARD-12X expressed sequence tag, GI10316320
XX
XX Sequence 1141 BP; 318 A; 351 C; 320 G; 152 T; 0 U; 0 Other;
XX
XX Query Match 26.4%; Score 110; DB 7; Length 1141;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-44;
XX Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ATGGGGGAGACGTGGCGAGGAGCTCCGACACTGACGGAGCTGAGAGAGACACTGTGG 60
XX DB 161 ATGGGGGAGACGTGGCGAGGAGCTCCGACACTGACGGAGCTGAGAGAGACACTGTGG 220
XX
XX QY 61 GAGATGATGAGAGAGCCACCGCAGATGATGATGATGATGATGATGATGATGATGATGAT 110
XX DB 221 GAGATGATGAGAGAGCCACCGCAGATGATGATGATGATGATGATGATGATGATGATGAT 270
XX
XX RESULT 6
XX ACA38053
XX ID ACA38053 standard; DNA; 534 BP.
XX AC ACA38053;
XX AC
XX DT 19-JUN-2003 (first entry)
XX DT
XX DE Prokaryotic essential gene #19710.
XX DE
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.
XX OS
XX Mycobacterium avium.
XX OS
XX PN WO20027183-A2.
XX PD
XX PF 03-OCT-2002.
XX PP 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zykkind JW;
XX PI Mail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR P-PSDB; ABU34183.
XX DR WPI; 2003-029926/02.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX PS Claim 14; SEQ ID NO 25923; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing of a gene in an operon required for
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 534 BP; 98 A; 187 C; 166 G; 83 T; 0 U; 0 Other;
XX
XX Query Match 4.8%; Score 20; DB 7; Length 534;
XX Best Local Similarity 100.0%; Pred. No. 9.6;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 305 TGGTCAACCGGGCTGACGCTT 324
XX DB 290 TGGTCAACCGGGCTGACGCTT 309

RESULT 7

ABN20134/c

ID ABN20134 standard; cDNA; 550 BP.

XX ABN20134;

DT 24-JUN-2002 (first entry)

XX Human ORFX polynucleotide sequence SEQ ID NO:8745.

XX Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypochloridism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.

XX Homo sapiens.

XX PN WO200192523-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US010836.

XX PK 30-MAY-2000; 2000US-0206132P.

XX PR 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach MD;

XX WPI; 2002-106308/14.

XX P-PSDB; ABP04382.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX PS Disclosure; SEQ ID NO 8745; 1037pp; English.

XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see table 1
 CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypochloridism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences

XX Sequence 550 BP; 177 A; 112 C; 137 G; 124 T; 0 U; 0 Other;

XX Query Match 4.6%; Score 19; DB 6; Length 550;

XX Best Local Similarity 100.0%; Pred. No. 30; Mismatches 19; Conservative 0; Indels 0; Gaps 0;

QY 48 GGAGCACTGTGGAGATG 66

DB 258 GGAGCACTGTGGAGATG 240

RESULT 8

ID AAA09358 standard; DNA; 867 BP.

XX AAA09358;

XX 10-AUG-2000 (first entry)

XX p75-NTR (neurotrophin receptor) extracellular portion DNA.

KW p75 neurotrophin receptor; extracellular; p75-NTR; death receptor;
 KW apoptosis; death signal domain; antagonist; cell survival; cytostatic;
 KW neuroprotective; antiparkinsonian; antidiabetic; anti-HIV; virucide;
 KW neurotrophic; anticonvulsant; cerebroprotective; ss.

XX Rattus sp.

XX Key Location/Qualifiers

XX Mat_peptide 115..867

XX WO200020578-A1.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-AU000860.

XX 06-OCT-1998; 98AU-00006353.

XX 07-OCT-1998; 98AU-00006351.

XX 01-JUN-1999; 99AU-00000701.

XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX Bartlett PF, Coulson EJ, Fieldew K, Baca M, Kilpatrick T;

XX Surindar C;

XX WPI; 2000-328933/28.

XX P-PSDB; AAY92372.

XX Claim 13; Page 71-72; 80pp; English.

XX This DNA encodes the extracellular portion (or ligand binding molecule)
 CC of rat p75 neurotrophin receptor (p75-NTR). p75-NTR is multifunctional
 CC and is capable of acting as a death receptor. Elevated p75-NTR expression
 CC results in increased death in vitro and in vivo. It has been determined
 CC that the death signal is not the cytoplasmic motif known as the death
 CC domain, but is a region adjacent to the membrane domain on p75-NTR.
 CC Identification of this region provides an opportunity to modulate cell
 CC survival by antagonizing the death signalling region or promoting
 CC apoptosis by providing cells with the genetic material to express the
 CC death signalling region adjacent, proximal, or otherwise juxtaposed or
 CC associated membrane of a cell or to express the death signalling region
 CC in multicentric form. The polypeptides and methods of the invention are
 CC useful for inhibiting, reducing or antagonizing p75-NTR-mediated death
 CC signals in neural cells. In particular, the methods are used for the
 CC treatment or prophylaxis of disease conditions associated with neural
 CC death or where cell death is to be promoted such as in treating or
 CC preventing cancer growth and/or development. Other diseases capable of
 CC treatment include neurodegenerative diseases, such as cerebral palsy,
 CC trauma induced paralysis, vascular ischemia associated with stroke,
 CC neural tumors, motorneuron diseases, Parkinson's disease, Huntington's
 CC disease, Alzheimer's disease, multiple sclerosis and peripheral
 CC neuropathies associated with diabetes, heavy metal or alcohol toxicity,

renal failure, and/or infectious diseases such as Herpes, rabella, measles, chicken pox, HIV and HTLV-1 (all claimed). The methods are also useful for treating neurons or glia damaged by trauma or disease (claimed). Animals which can be treated include humans, livestock CC animals, laboratory test animals, companion animals, and captive wild CC animals

XX
SQ Sequence 867 BP, 177 A; 260 C; 288 G; 142 T; 0 U; 0 Other;

Query Match 4.6%; Score 19; DB 3; Length 867;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 388 AGGAACCTCAGGCTCTCG 406
|||||
DB 771 AGGAACCTCAGGCTCTCG 753

RESULT 9
ADBS8069/c
ID ADBS8069 standard; DNA; 3259 BP.
XX
AC ADBS8069;
XX
DT 04-DEC-2003 (first entry)
XX
DE Toxicity-related gene, SEQ ID 3095.
XX
KM Toxic; toxin; gene expression profile; hepatotoxicity; liver;
XX drug screening; toxicity assay; ds.
XX
OS Unidentified.
XX
PN WO2003064624-A2.
XX
PD 07-AUG-2003.
XX
PF 31-JAN-2003; 2003WO-US003194.
XX
PR 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Casle A, Elashoff M;
XX
DR WPI; 2003-689530/65.
XX
XX
PT Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
PT the compound.
XX
PS Claim 1; SEQ ID NO 3095; 1156pp; English.
XX
XX
XX The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

```

CC      at ftp.wjpo.int/pub/published_pcr_sequences.
XX
SQ      Sequence 3259 BP; 717 A; 976 C; 913 G; 653 T; 0 U; 0 Other;
XX
Query Match          4.6%; Score 19; DB 9; Length 3259;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      388 AGAAGCTCAGGCTCTGG 406
      |||||
DB      770 AGGAAGCTCAGGCTCTGG 752

RESULT 10
AAA09356/c
ID      AAA09356 standard; cDNA; 3260 BP.
XX
AC      AAA09356;
XX
DT      10-AUG-2000 (first entry)
XX
DE      p75-NTR (neurotrophin receptor) coding sequence.
XX
KW      p75 neurotrophin receptor; p75-NTR; death receptor; apoptosis;
KM      death signal domain; antagonist; cell survival; cytostatic;
KM      neuroprotective; antiparkinsonian; antidiabetic; anti-HIV; virucide;
XX      neurotrophic; anticonvulsant; cerebroprotective; ss.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      115..1392
              /*tag= a
PN      WO200020578-A1.
XX
FD      13-APR-2000.
XX
PF      05-OCT-1999; 99WO-AU000860.
XX
PR      06-OCT-1998; 98AU-00006353.
PR      07-OCT-1998; 98AU-00006351.
PR      01-JUN-1999; 99AU-00000701.
XX
PA      (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI      Battlert PF, Coulson EJ, Fieldew K, Baca M, Kilpatrick T,
PI      Surindar C;
XX
XX      WP1; 2000-328933/28.
DR      P-PSDB; AAY923370.
XX
PT      Novel nucleic acids encoding a peptide capable of facilitating the death
XX      of a cell, useful for antagonizing cell death signal function and
XX      promoting cell death, e.g. for treating cancer.
PS      Claim 9; Page 65-69; 80pp; English.
XX
XX      This CDNA encodes human p75 neurotrophin receptor (p75-NTR). p75-NTR is
CC      multifunctional and is capable of acting as a death receptor. Elevated
CC      p75-NTR expression results in increased death in vitro and in vivo. It
CC      has been determined that the death signal is not the cytoplasmic motif
CC      known as the death domain, but is a region adjacent the membrane domain
CC      on p75-NTR. Identification of this region provides an opportunity to
CC      modulate cell survival by antagonising the death signalling region or
CC      promoting apoptosis by providing cells with the genetic material to
CC      express the death signalling region adjacent, proximal, or otherwise
CC      juxtaposed or associated membrane of a cell or to express the death
CC      signalling region in multicentric form. The polypeptides and methods of the
CC      invention are useful for inhibiting, reducing or antagonizing p75-NTR-
CC      mediated death signals in neural cells. In particular, the methods are
CC      used for the treatment or prophylaxis of disease conditions associated
CC      with neural death or where cell death is to be promoted such as in

```

CC treating or preventing cancer growth and/or development. Other diseases
CC capable of treatment include neurodegenerative diseases, such as cerebral
CC palsy, trauma induced paralysis, vascular ischemia associated with
CC stroke, neural tumors, motorneuron diseases, Parkinson's disease,
CC Huntington's disease, Alzheimer's disease, multiple sclerosis and
CC peripheral neuropathies associated with diabetes, heavy metal or alcohol
CC toxicity, renal failure, and/or infectious diseases such as Herpes,
CC rubella, measles, chicken pox, HIV and HTLV-1 (all claimed). The methods
CC are also useful for treating neurons or glia damaged by trauma or disease
CC (claimed). Animals which can be treated include humans, livestock
CC animals, laboratory test animals, companion animals, and captive wild
CC animals

XX Sequence 3260 BP; 718 A; 976 C; 913 G; 653 T; 0 U; 0 Other;
SQ

Query Match 4.6%; Score 19; DB 3; Length 3260;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 AGGAACCTCAGGCTCCTGG 406
Db 771 AGGAACCTCAGGCTCCTGG 753

RESULT 11
AAS98203
ID AAS98203 standard; cDNA; 3948 BP.
XX AAS98203;
AC
DT 12-MAR-2002 (first entry)
XX
XX DNA encoding plakoglobin interacting protein #3.
XX
XX Human; plakoglobin; cytoskeletal; osteopathic; dermatological; cardiac;
XX plakoglobin related disease; skin carcinoma; acantholytic disease;
XX basal cell carcinoma; squamous cell carcinoma; Naxos disease; PCR primer;
XX extramammary Paget's disease; heart disease; skin blistering;
XX subcorneal acantholysis; Grover's disease; Halley-Halley's disease;
XX Darier's disease; ectodermal dysplasia; skin fragility syndrome; ss.
XX
XX Homo sapiens.
XX
XX MO200185933-A2.
XX
XX 15-NOV-2001.
XX
XX 02-MAY-2001; 2001MO-EP04872.
XX
XX 09-MAY-2000; 2000EP-00201668.
XX
XX (VLAAs-) VLAAS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Van Roy F, Bonne S, Vanlandschoot A;
XX
XX WPI; 2002-0632246/08.
XX
XX P-PSDB; AAU73247.
XX
XX New polypeptide, useful for treating skin carcinoma or acantholytic
XX disease such as Grover's and Darier's disease, comprises a protein
XX interacting with human plakoglobin and involved in transduction of
XX plakoglobin related signal to nucleus.
XX
XX Claim 7; Fig 3; 98pp; English.
XX
XX The invention relates to an isolated plakoglobin interacting polypeptide
XX (I). (I) is useful as a medicament and in the manufacture of a medicament
XX for treating plakoglobin related diseases, such as skin carcinoma or an
XX acantholytic disease, and to screen compounds that interfere with the
XX interaction of the polypeptide with plakoglobin. The plakoglobin related
XX diseases include basal cell carcinoma, squamous cell carcinoma,
XX extramammary Paget's disease, Naxos disease, heart diseases, skin
XX blistering and acantholytic diseases such as subcorneal acantholysis,

CC Grover's disease, Halley-Halley's disease or Darier's disease, and
CC ectodermal dysplasia/skin fragility syndrome. AAS98201- AAS98298
CC represent novel human plakoglobin interacting protein coding sequences
CC and PCR primers of the invention

XX Sequence 3948 BP; 717 A; 1218 C; 1327 G; 686 T; 0 U; 0 Other;
SQ

Query Match 4.6%; Score 19; DB 6; Length 3948;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 GGACGAGGAGGAGGCTGCTG 174
Db 224 GGACGAGGAGGAGGCTGCTG 242

RESULT 12
AAS05388
ID AAS05388 standard; cDNA; 3949 BP.
XX AAS05388;
AC
DT 12-SEP-2001 (first entry)
XX
XX Human caspase recruitment domain, CARD-10 cDNA sequence.
XX
XX Human; caspase recruitment domain; CARD-10; Bcl-10; NF-kappaB; apoptosis;
XX hyperproliferative disorder; autoimmune; neurological;
XX inflammatory disorder; viral infection; stress-related response; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 41..3139
XX /*tag= a
XX /product= "CARD-10"
XX /note= "ORF is specifically claimed"
XX
XX MO200140468-A2.
XX
XX 07-JUN-2001.
XX
XX 01-DEC-2000; 2000WO-US032716.
XX
XX 03-DEC-1999; 99US-0168780P.
XX 18-FEB-2000; 2000US-00507533.
XX 25-FEB-2000; 2000US-00513904.
XX 10-OCT-2000; 2000US-00685791.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Bertin J;
XX
XX WPI; 2001-367809/38.
XX
XX P-PSDB; AAU01206.
XX
XX Novel caspase recruitment domain (CARD) proteins, CARD-9, CARD-10, CARD-
XX 11, useful as targets for therapy, as immunogens, and in screening and
XX detection assays.
XX
XX Claim 2; Fig 10A-10C; 145pp; English.
XX
XX The present sequence encoding for novel human caspase recruitment domain,
XX CARD-10 is isolated from a human skin cDNA library. Also described are
XX novel human sequences for CARD-9 and CARD-11 (AAU01205, AAU01207) and rat
XX CARD-9 (AAU01204). CARD-9, CARD-10 and CARD-11 interact with Bcl-10 which
XX is thought to activate nuclear factor (NF)-kappaB and apoptosis. The
XX sequences of the invention can be used for treating a disorder associated
XX with abnormal levels of apoptosis by modulating the expression or
XX activity of CARD-9, CARD-10, or CARD-11. They can be used for the
XX treatment of hyperproliferative disorders (e.g. cancer), autoimmune
XX disorders (e.g. systemic lupus erythematosus), neurological disorders
XX (e.g. Alzheimer's disease), inflammatory disorders (e.g. Crohn's

Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 TCTACACCTGTGTCACCGG 314
Db 1307 TCTACACCTGTGTCACCGG 1289

Search completed: March 21, 2004, 23:44:55
Job time : 261 secs

RESULT 15
AAZ22282/c
ID AAZ22282 standard; DNA; 5790 BP.

XX AAZ22282;

XX 06-DEC-1999 (first entry)

XX Nucleic acid sequence of KIAA380.

XX RGS; guanine nucleotide exchange factor; GEF; RGS domain; GAs; cancer;
KM G protein alpha subunit; cell proliferation; growth control; hemostasis;
KM morphogenesis; stress fiber formation; integrin-mediated interaction;
KM embryonic development; tumor cell growth; cell death; leukocyte homing;
KM bone resorption; clot retraction; db1 homology domain; mechanical stress;
KM plectstrin homology domain; KIAA380; ss.

XX Homo sapiens.

XX WO9947557-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99MO-US006051.

XX 18-MAR-1998; 98US-0078634P.

XX (ONYX-) ONYX PHARM INC.

XX Boliag G, Hart MJ, Roscoe W, Polakis P, Sternweis P, Kozasa T;
PI Jiang X;

XX WPI; 1999-571821/48.

XX P-PSDB; AAY41010.

PT New isolated RGS-GEF polypeptides, used to develop products for
modulating, e.g. cell proliferation and integrin-mediated interactions.

PS Example; Fig 13; 75pp; English.

XX The invention relates to isolated RGS-guanine nucleotide exchange factor
(GEF) polypeptides. The novel isolated RGS-GEF polypeptide consists of an
RGS domain of a GEF protein and does not comprise a db1 homology (DH)
domain or a plectstrin homology (PH) domain. The RGS-GEF polypeptides can
be used for modulating an activity of a G protein alpha subunit (GAs).
XX The products can be used for the regulation of biological pathways in
which a RGS-GEF polypeptide is involved, particularly pathological
conditions, e.g. cell proliferation (e.g. cancer), morphogenesis, growth
control, stress fiber formation, and integrin-mediated interactions, such
as embryonic development, tumor cell growth and metastasis, programmed
cell death, hemostasis, leukocyte homing and activation, bone resorption,
clot retraction, and the response of cells to mechanical stress. The
XX products can also be used for detection, diagnosis and production of
transgenic animals. The present sequence represents the nucleic acid
sequence of KIAA380

XX Sequence 5790 BP; 1415 A; 1640 C; 1515 G; 1220 T; 0 U; 0 Other;

Query Match 4.6%; Score 19; DB 2; Length 5790;

Best Local Similarity 100.0%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 TCTACACCTGTGTCACCGG 314
Db 1596 TCTACACCTGTGTCACCGG 1578

COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Filting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-276-852-35

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGGAGAGCCACCGCCAC 84
DB 31 ATGGAGAGCCACCGCCAC 14

RESULT 3
US-08-133-011-107/c
Sequence 107, Application US/08133011
Patent No. 5658727
GENERAL INFORMATION:
APPLICANT: Kang, Andray
APPLICANT: Barbas, Carlos
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
TITLE OF INVENTION: PHAGEMIDS
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5658727th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/133,011
FILING DATE: 08-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/683,602
FILING DATE: 10-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Filting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-133-011-107

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGGAGAGCCACCGCCAC 84
DB 31 ATGGAGAGCCACCGCCAC 14

RESULT 4
US-08-322-730A-107/c
Sequence 107, Application US/08322730A
Patent No. 5759817
GENERAL INFORMATION:
APPLICANT: Barbas, Carlos
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
TITLE OF INVENTION: PHAGEMIDS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10550 No. 5759817th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,730A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
APPLICATION NUMBER: US 07/683,602
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bingham, Douglas A
REGISTRATION NUMBER: 32,457
REFERENCE/DOCKET NUMBER: SCRO707P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-784-2937

TELEFAX: 619-784-9399
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-322-730A-107

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 5

US-08-387-874-79/c
Sequence 79, Application US/08387874
Patent No. 5770356

GENERAL INFORMATION:

APPLICANT: Light, Paul L., II
APPLICANT: Lerner, Richard A.

TITLE OF INVENTION: PHAGMIDS COEXPRESSING A SURFACE
NUMBER OF INVENTION: RECEPTOR AND A SURFACE HETEROLOGOUS PROTEIN

NUMBER OF SEQUENCES: 97

CORRESPONDENCE ADDRESS:
ADDRESS: The Scripps Research Institute, Office of

STREET: 10666 No. 5770356th Torrey Pines Road, Suite 220,
STREET: Mail Drop 1PC8

CITY: La Jolla
STATE: CA

COUNTRY: USA

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/387,874
FILING DATE: 22-FEB-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/08364
FILING DATE: 03-SEP-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/941,369

FILING DATE: 04-SEP-1992
ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 303.1
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:

LENGTH: 48 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO

ANTI-SENSE: NO
US-08-387-874-79

Query Match 4.3%; Score 18; DB 1; Length 48;

Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 6

US-08-899-575-35/c
Sequence 35, Application US/08899575
Patent No. 5770440

GENERAL INFORMATION:

APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F

APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170

CORRESPONDENCE ADDRESS:
ADDRESS: The Scripps Research Institute, Office of

STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
STREET: Mail Drop 1PC8

CITY: La Jolla
STATE: CA

COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994

APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148

FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:

LENGTH: 48 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO

ANTI-SENSE: NO
US-08-899-575-35

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
|||||
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 7

US-08-899-575-35/c
Sequence 35, Application US/08899575
Patent No. 5804440
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-899-575-35

Query Match 4.3%; Score 18; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 8
US-08-383-619-107/c
Sequence 107, Application US/08383619
Patent No. 5955341
GENERAL INFORMATION:
APPLICANT: Kang, Angray
APPLICANT: Barbas, Carlos
APPLICANT: Lerner, Richard
TITLE OF INVENTION: HETEROLOGOUS RECEPTOR LIBRARIES USING
TITLE OF INVENTION: PHAGMIDS

NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: DOUGLAS A. BINGHAM
STREET: 11300 Sorrento Valley Road, Suite 200
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,619
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,680
FILING DATE:
APPLICATION NUMBER: US/07/683,602
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bingham, Douglas A.
REGISTRATION NUMBER: 32,457
REFERENCE/DOCKET NUMBER: SCR0371P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-1555
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-383-619-107

Query Match 4.3%; Score 18; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 9
US-08-888-366-36/c
Sequence 36, Application US/08888366
Patent No. 5972656
GENERAL INFORMATION:
APPLICANT: Lopez, Osvaldo
APPLICANT: Wylie, Fred W.
APPLICANT: Wagner, Fred W.
TITLE OF INVENTION: Mercury Binding Polypeptides and Nucleotides Coding Therefore
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 90 South 7th Street, 3100 No. 5972656west Ctr.
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/888,366
FILING DATE: 03-JUL-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/187,407
FILING DATE: 27-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,542
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/493,299
FILING DATE: 14-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/324,392
FILING DATE: 14-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.39USC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-888-366-36

Query Match 4.3%; Score 18; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGGAGAGCCACCGCCAC 84
DB 31 ATGGAGAGCCACCGCCAC 14

RESULT 10
US-08-767-128-43/C
Sequence 43, Application US/08767128
Patent No. 6111079
GENERAL INFORMATION:
APPLICANT: WYLLIE, DANE E.
APPLICANT: LOPEZ, OSVALDO
APPLICANT: MORRAY, PETER JOSEPH
APPLICANT: GOEBEL, PETER
TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSER: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6111079west Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,128
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 04-DEC-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09258
FILING DATE: 05-JUN-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/541,373
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,798
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carter, Charles G.
REGISTRATION NUMBER: 35,093
REFERENCE/DOCKET NUMBER: 8648.49USF1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/371-5278
TELEFAX: 612/332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-767-128-43

Query Match 4.3%; Score 18; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGGAGAGCCACCGCCAC 84
DB 31 ATGGAGAGCCACCGCCAC 14

RESULT 11
US-08-907-739-107/C
Sequence 107, Application US/08907739
Patent No. 6235469
GENERAL INFORMATION:
APPLICANT: Kang, Angray
APPLICANT: Barbos, Carlos
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSER: The Scripps Research Institute, Office of
STREET: 10666 No. 6235469th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/907,739
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/133,011
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-907-739-107

Query Match 4.3%; Score 18; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 12
US-08-972-564-19/c
Sequence 19, Application US/08972564
Patent No. 6376170
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R.
APPLICANT: Williamson, R. Anthony
APPLICANT: Sanna, Pietro Paolo
TITLE OF INVENTION: LIGAND CAPTURE-DIRECTED SELECTION OF ANTIBODY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,564
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/316,914
FILING DATE: October 3, 1994
ATTORNEY/AGENT INFORMATION:
NAME: LISA A. HAILE, P.H.D.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD-3809
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS

LOCATION: 1..48
US-08-972-564-19

Query Match 4.3%; Score 18; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 13
US-09-729-597-107/c
Sequence 107, Application US/09729597
Patent No. 6468738
GENERAL INFORMATION:
APPLICANT: Kang, Angray
Barbas, Carlos
Lerner, Richard A.
TITLE OF INVENTION: HETERODIMERIC RECEPTOR LIBRARIES USING PHAGMIDS
NUMBER OF SEQUENCES: 161
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of Patent Counsel
STREET: 10666 No. 6468738th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/729,597
FILING DATE: 04-Dec-2000
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/133,011
FILING DATE: 1994-09-29
APPLICATION NUMBER: US 07/826,623
FILING DATE: 27-JAN-1992
APPLICATION NUMBER: PCT/US 92/03091
FILING DATE: 10-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRF 238.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 107:
US-09-729-597-107

Query Match 4.3%; Score 18; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
Db 31 ATGAGAGCCACCGCCAC 14

RESULT 14
PCT-US93-08364-79/c
Sequence 79, Application PC/TUS9308364
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PHAGEMIDS COEXPRESSION A SURFACE
RECEPTOR AND A SURFACE HETEROLOGOUS PROTEIN
NUMBER OF SEQUENCES: 97
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08364
FILING DATE: 03-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/941,369
FILING DATE: 04-SEP-1992
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US93-08364-79

Query Match 4.3%; Score 18; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

RESULT 15
PCT-US95-00067-21/c
Sequence 21, Application PC/TUS9500067
GENERAL INFORMATION:
APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO HERPES
SIMPLEX VIRUS AND METHODS THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Judas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00067
FILING DATE: 04-JAN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE//DOCKET NUMBER: PD-3229
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..48
PCT-US95-00067-21

Query Match 4.3%; Score 18; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGAGCCACCGCCAC 84
DB 31 ATGAGAGCCACCGCCAC 14

Search completed: March 22, 2004, 03:19:52
Job time : 81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2004, 23:45:00 ; Search time 3164 Seconds
(without alignment)
3935.692 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417
Sequence: 1 atgggggaactgtccgcag.....ggctcctgtaaccccgagm 417

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estbta:*
2: em_estbta:*
3: em_estbta:*
4: em_estbta:*
5: em_estbta:*
6: em_estbta:*
7: em_estbta:*
8: em_estbta:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estbta:*
16: em_estbta:*
17: em_estbta:*
18: em_estbta:*
19: em_estbta:*
20: em_estbta:*
21: em_estbta:*
22: em_estbta:*
23: em_estbta:*
24: em_estbta:*
25: em_estbta:*
26: em_estbta:*
27: em_estbta:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	351	84.2	931	13	BU956261
2	349	83.7	3015	29	AY414357 Homo sapi
3	338	81.1	3015	29	AY414358 Pan trogl
4	110	26.4	1141	10	BE867544

5	47	11.3	549	10	BF079819	BF079819	230513	MA
C	33	7.9	806	12	BF031835	BF031835	UI-1-CFO-	
7	33	7.9	806	10	BF100075	BF100075	601752245	
8	33	7.9	1600	11	AK086176	AK086176	Mus muscu	
9	23	5.5	2996	29	AY414359	AY414359	Mus muscu	
10	23	5.5	601	10	BE291786	BE291786	601084552	
C	11	5.0	478	28	AO850034	AO850034	IMAFV1.1	
12	21	5.0	514	12	BC090138	BC090138	ut57P02.Y	
13	21	5.0	537	29	L80078Y	L80078Y	AL334405	
C	14	5.0	576	28	A2629378	A2629378	2M0107F06	
15	21	5.0	655	12	BI649379	BI649379	603277822	
16	21	5.0	1024	13	BY702884	BY702884	BY702884	
17	21	5.0	1075	11	AK002346	AK002346	Mus muscu	
C	20	4.6	1014	29	AY402215	AY402215	Mus muscu	
19	20	4.6	234	14	W83246	W83246	mi27F1.1	
20	19	4.6	239	9	AI622099	AI622099	486032G10	
C	21	4.6	239	9	AV122681	AV122681	AV122681	
22	19	4.6	360	9	AV193176	AV193176	AV193176	
23	19	4.6	370	10	BF018592	BF018592	ux7961.1	
24	19	4.6	371	28	A2615770	A2615770	IM0445H13	
25	19	4.6	381	12	BM076573	BM076573	TEEST-A00	
26	19	4.6	392	12	BG717357	BG717357	602689691	
C	27	4.6	409	10	BF507352	BF507352	UI-H-BM1-	
28	19	4.6	431	12	BJ178531	BJ178531	BJ178531	
C	29	4.6	433	12	BI740905	BI740905	GC91D07.Y	
30	19	4.6	438	14	CA389497	CA389497	CS09H05.Y	
31	19	4.6	444	10	AM829774	AM829774	ra44C08.Y	
C	32	4.6	446	9	AA647911	AA647911	vm08F05.X	
33	19	4.6	453	12	BI740621	BI740621	gc83811.Y	
34	19	4.6	453	12	BM001080	BM001080	1031093A1	
C	35	4.6	471	29	CE605544	CE605544	1191-988-	
36	19	4.6	484	14	CB727438	CB727438	AMGNNUC.M	
C	37	4.6	484	12	BJ184852	BJ184852	BJ184852	
38	19	4.6	486	10	BF114874	BF114874	7187H01.X	
C	39	4.6	515	28	BH226597	BH226597	100613C00	
40	19	4.6	519	12	BJ196726	BJ196726	BJ196726	
C	41	4.6	519	12	BJ197185	BJ197185	BJ197185	
42	19	4.6	519	12	BM877430	BM877430	1494H09.Y	
43	19	4.6	527	28	AF106737	AF106737	AF106737	
C	44	4.6	537	12	BQ039741	BQ039741	gd02605.Y	
45	19	4.6	537	12	BQ039741	BQ039741	gd02605.Y	

ALIGNMENTS

RESULT 1
BU956261
LOCUS
DEFINITION AGENCOURT 10613482 NIH_MGC_107 Homo sapiens CDNA clone
IMAGE:6729809 5', mRNA sequence.
ACCESSION BU956261
VERSION BU956261.1 GI:24185833
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
JOURNAL
COMMENT
BU956261 931 bp mRNA linear EST 21-OCT-2002
AGENCOURT 10613482 NIH_MGC_107 Homo sapiens CDNA clone
IMAGE:6729809 5', mRNA sequence.
BU956261.1 GI:24185833
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 931)
NIH-MGC http://mgs.nci.nih.gov/
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM3053 row: 5 column: 16
High quality sequence stop: 594.

COMMENT
This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
location/Qualifiers
1.3015
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
1.3015
/gene="CARD14"
/locus_tag="HCM5197"

ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 4.2e-146;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

81.1%; Score 338; DB 29; Length 3015;
1 ATGGGGGAAGTGTGCGCCAGAGGACTCCGCACTGACGCACTGAGAGAGACACTGTGG 60
1 ATGGGGGAAGTGTGCGCCAGAGGACTCCGCACTGACGCACTGAGAGAGACACTGTGG 60
61 GAGATGATGAGAGAGCCAGCCGCAAGATGTACGCTGCACTTGCCTCCAGCCGCTCACC 120
61 GAGATGATGAGAGAGCCAGCCGCAAGATGTACGCTGCACTTGCCTCCAGCCGCTCACC 120
121 CCTCACTGGCGGCGGAGGAGGAGTGTGCTGCACTGAGAGAGAGAGGAGTGTGCACTG 180
121 CCTCACTGGCGGCGGAGGAGGAGTGTGCTGCACTGAGAGAGAGAGGAGTGTGCACTG 180
181 CCGCGGCTGCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
181 CCGCGGCTGCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
241 GGGAGAGAGCGGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
241 GGGAGAGAGCGGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
301 ACCCTGGTCAACGGGGCTGACGCTGATGTGACTTACG 338
301 ACCCTGGTCAACGGGGCTGACGCTGATGTGACTTACG 338

RESULT 4
LOCUS
DEFINITION
VERSION
KEYWORDS
SOURCE
ORGANISM
1141 bp mRNA linear EST 20-OCT-2000
601443043P1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3847282 5',
mRNA sequence.
BE867544
BE867544
BE867544.1 GI:10316320
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1141)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LIA9561 row: e column: 11
High quality sequence start: 4
High quality sequence stop: 269.
Location/Qualifiers
1.1141
/organism="Homo sapiens"

FEATURES
source
1.1141
/organism="Homo sapiens"

ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 3.9e-40;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

26.4%; Score 110; DB 10; Length 1141;
1 ATGGGGGAAGTGTGCGCCAGAGGACTCCGCACTGACGCACTGAGAGAGACACTGTGG 60
1 ATGGGGGAAGTGTGCGCCAGAGGACTCCGCACTGACGCACTGAGAGAGACACTGTGG 60
61 GAGATGATGAGAGAGCCAGCCGCAAGATGTACGCTGCACTTGCCTCCAGCCGCTCACC 110
61 GAGATGATGAGAGAGCCAGCCGCAAGATGTACGCTGCACTTGCCTCCAGCCGCTCACC 110
221 GAGATGATGAGAGAGCCAGCCGCAAGATGTACGCTGCACTTGCCTCCAGCCGCTCACC 270

RESULT 5
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
549 bp mRNA linear EST 18-OCT-2000
BF079819
BF079819
BF079819
BF079819.1 GI:10873649
EST.
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
1 (bases 1 to 549)
Fahnenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Perrea, G., Sultana, R.,
Quackenbush, J. and Keel, J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smitht@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCCACTCAGACG
Plate: 47 row: M column: 12
Seq primer: ATTATGCTGACCTATAG.
Location/Qualifiers
1.549
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9623"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

FEATURES
source
1.549
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9623"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

ORIGIN
Query Match
11.3%; Score 47; DB 10; Length 549;

Best Local Similarity 100.0%; Pred. No. 6.7e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AAGACGGGGCCATCGCTTCTGAGAGCCCTGAGTTCACACCC 290
|||||
314 AAGACGGGGCCATCGCTTCTGAGAGCCCTGAGTTCACACCC 360
|||||

RESULT 6
BQ031835 492 bp mRNA linear EST 27-MAR-2002
LOCUS UI-1-CFO-ach-c-03-0-UI 3' mRNA sequence.
DEFINITION UI-1-CFO-ach-c-03-0-UI 3', mRNA sequence.

ACCESSION BQ031835
VERSION BQ031835.1 GI:19767114
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 492)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Minoru Ko, Dr. Janet Rossent, Dr. Janet

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-49, >AT-richLow_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source Location/Qualifiers

1..492
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-1-CFO-ach-c-03-0-UI"
/tissue_type="Trophoblast"
/lab_host="DH10B (Life Technologies)"
/clone_1lb="NCI CGAP Pltr1"
/note="Organ: Placenta; Vector: pRTT3-Pac (Pharmacia) with
a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP Pltr1 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pRTT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are GTGG, ATCAT, GGATG,
GGTGG. For additional information, contact: Bento Soares,
bento-soares@uiowa.edu
TAG_11505=Placenta mouse 11.5-12.5 days
TAG_11506=ATCAT"
TAG_SEQ=ATCAT"

Query Match. 7.9%; Score 33; DB 12; Length 492;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 CTCACCCCTACTGCGCCAGGCCAAGGTGCTG 147
|||||

DB 316 CTCACCCCTACTGCGCCAGGCCAAGGTGCTG 284

RESULT 7
BF100075 806 bp mRNA linear EST 19-OCT-2000
LOCUS 601752245F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3979883 5',
DEFINITION mRNA sequence.

ACCESSION BF100075
VERSION BF100075.1 GI:10882601
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 806)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomix, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM9174 row: j column: 12
High quality sequence, stop: 668.
Location/Qualifiers

1..806
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979883"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Mam1"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match. 7.9%; Score 33; DB 10; Length 806;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 CTCACCCCTACTGCGCCAGGCCAAGGTGCTG 147
|||||
DB 418 CTCACCCCTACTGCGCCAGGCCAAGGTGCTG 450
|||||

RESULT 8
AK086176 1600 bp mRNA linear HTC 20-SEP-2003
LOCUS AK086176
DEFINITION Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched
library, clone: D93010J19 product: CASPASE RECRUITMENT DOMAIN
PROTEIN 14 (BCL10-INTERACTING MAGUK PROTEIN 2) (BIM2) [Mus
musculus], full insert sequence.

ACCESSION AK086176
VERSION AK086176.1 GI:26103277
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hoshizume, W., Hayashida, K., Hayatsu, N., Hiramatsu, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 1600
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D93001019"
/db_xref="MGI:2423547"
/db_xref="taxon:10090"
/clone="D93001019"

misc_feature
/tissue type="head"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev stage="15 days embryo"
1. 1600
/note="CASBASE RECRUITMENT DOMAIN PROTEIN 14 (BC10-INTERACTING MAGUK PROTEIN 2) (BIMP2) [Mus musculus] (SWISSPROT|Q99KF0, evidence: FASTA, 85.1%id, 76.9%length, match=2706 and 100% to the mouse sequence)"
ORIGIN
Query Match 7.9%, Score 33, DB 11, Length 1600;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 33, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
Qy 115 CTCACCCCTCCTGCGCCAGGCCAAGTGTGCTG 147
Db 688 CTCACCCCTCCTGCGCCAGGCCAAGTGTGCTG 720
RESULT 9
AY143359
LOCUS AY143359 2996 bp DNA linear GSS 17-DEC-2003
DEFINITION Mus musculus CARD14 gene, VIRUTAL TRANSCRIPT, partial sequence.
ACCESSION AY143359
VERSION AY143359.1 GI:39770321
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 2996)
REFERENCE 1
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Smutsy, J.J., Adams, M.D., and Cargill, M.
TITLE Interfering nonneutral evolution from human-chimp-mouse orthologous gene titles
JOURNAL Science 302 (5652), 1960-1963 (2003)
MEDLINE 14671302
PUBMED 14671302
REFERENCE 2 (bases 1 to 2996)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Smutsy, J.J., Adams, M.D., and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
Location/Qualifiers
1. 2996
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/gene="CARD14"
/locus_tag="HCM5197"
ORIGIN
Query Match 7.9%, Score 33, DB 29, Length 2996;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 33, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
Qy 115 CTCACCCCTCCTGCGCCAGGCCAAGTGTGCTG 147
Db 115 CTCACCCCTCCTGCGCCAGGCCAAGTGTGCTG 147
RESULT 10
BE291786 601 bp mRNA linear EST 13-JUL-2000
LOCUS BE291786
DEFINITION 601084552P1 NCI_GAP_Mam6 Mus musculus cDNA clone IMAGE:3498685 5',

ACCESSION BE291786 mRNA sequence.
 VERSION BE291786.1 GI:9173747
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 601)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@b-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1M8554 row: h column: 14
 High quality sequence stop: 582.
 Location/Qualifiers
 1..601
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3498685"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="MDH10B"
 /clone_lib="NCI-CGAP_Mame"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt: Site: 2; NCI: Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

ORIGIN
 Query Match 5.5%; Score 23; DB 10; Length 601;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 AGCCTGAAGTTCACACCTGCA 293
 DB 10 AGCCTGAAGTTCACACCTGCA 32

RESULT 11 478 bp DNA linear GSS 25-MAY-2001
 LOCUS A0850034
 DEFINITION LMAJFV1 lms3a09.v1 Leishmania major FV1 random genomic library
 Leishmania major genomic clone LMAJFV1 lms3a09 5' similar to
 TR:92188 P92188 PROTEASOME 29 KD SUBUNIT 1, genomic survey
 sequence.
 A0850034
 A0850034.1 GI:6054682
 GSS.
 Leishmania major
 Leishmania major
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 1 (bases 1 to 478)
 Akopyants, N.S., Clifton, S.W., Martin, J., Page, D., Wylie, T., Li, L., Kiehl, J.C., Roos, D.S. and Beverley, S.M.
 A survey of the Leishmania major Friedlin strain V1 genome by shotgun sequencing: a resource for DNA microarrays and expression profiling
 JOURNAL Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)
 MEDLINE 21192569
 PUBMED 11295190

ACCESSION A0850034
 VERSION A0850034.1 GI:6054682
 KEYWORDS GSS.
 SOURCE Leishmania major
 ORGANISM Leishmania major
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 1 (bases 1 to 478)
 Akopyants, N.S., Clifton, S.W., Martin, J., Page, D., Wylie, T., Li, L., Kiehl, J.C., Roos, D.S. and Beverley, S.M.
 A survey of the Leishmania major Friedlin strain V1 genome by shotgun sequencing: a resource for DNA microarrays and expression profiling
 JOURNAL Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)
 MEDLINE 21192569
 PUBMED 11295190

COMMENT Other GSSs: lms3a09.x1
 Contact: Akopyants, NS / Beverley, SM
 Washn Leishmania Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 Library construction: Natalia S. Akopyants, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 If using this information please cite:
 N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major Friedlin strain V1 genome by shotgun sequencing' and the Washington University Genome Sequencing Center for information on obtaining clone material please contact: Natalia S. Akopyants Ph.D. (nataliab@orcim.wustl.edu) and/or Stephen M. Beverley Ph.D. (beverley@orcim.wustl.edu)
 Seq primer: -40RP from Gibco
 Class: shotgun
 High quality sequence stop: 369.
 Location/Qualifiers
 1..478
 /organism="Leishmania major"
 /mol_type="genomic DNA"
 /strain="Friedlin strain V1"
 /db_xref="taxon:5664"
 /clone="LMAJFV1 lms3a09"
 /lab_host="TOP10 (Invitrogen)"
 /clone_lib="Leishmania major FV1 random genomic library"
 /note="Vector: pZero-2 (Invitrogen); Site: 1; EcoRV; Genomic DNA was isolated from stationary phase cells. For this library, DNA was sheared to give a tight size distribution of 1-1.5kb fragments, blunt-ended with T4 DNA polymerase, dephosphorylated with Shrimp Alkaline Phosphatase and ligated into pZero-2 vector's EcoRV site."

ORIGIN
 Query Match 5.0%; Score 21; DB 28; Length 478;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 GGACGAGAGAGAGTGTGCA 176
 DB 149 GGACGAGAGAGAGTGTGCA 129

RESULT 12 514 bp mRNA linear EST 26-JAN-2001
 LOCUS BG090138
 DEFINITION U57B02.V1 Soares mouse NMB bccl1 Mus musculus cDNA clone
 IMAGE:333189 5', mRNA sequence.
 BG090138
 BG090138.1 GI:12572701
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 514)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@b-remail.nih.gov
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 NCI:1076063
 Seq primer: -40RP from Gibco
 High quality sequence stop: 475.
 Location/Qualifiers
 1..514
 /organism="Mus musculus"

ACCESSION BG090138
 VERSION BG090138.1 GI:12572701
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 514)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@b-remail.nih.gov
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 NCI:1076063
 Seq primer: -40RP from Gibco
 High quality sequence stop: 475.
 Location/Qualifiers
 1..514
 /organism="Mus musculus"

JOURNAL
 COMMENT
 TITLE
 AUTHORS
 REFERENCE
 TITLE
 JOURNAL
 MEDLINE
 PUBMED

```

/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:333189"
/lab_host="DH10B (phage-resistant)"
/clone_lib="scars_mouse_NX9B_bcell"
/notes="Organ: germinal B-cell; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5].
TGTTACCAATCTGAAGGAGGAGCGCCGCTGTTTCTTTTCTTTTCTTTTCTTTT
T 3'); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized; constructed by Benito Soares and
M.Fatima Bonaldi."

```

ORIGIN

```

Query Match      5.0%; Score 21; DB 12; Length 514;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 115 CTCACCCCTACCTGCGCCAG 135

DB 335 CTCACCCCTACCTGCGCCAG 355

```

RESULT 13
L8007BY      537 bp  DNA      linear  GSS 29-APR-2000
LOCUS      Leishmania major Friedlin cosmid L8007.2 t7 end-sequence, similar
DEFINITION to T8R131148 Trypanosoma brucei rhodesiense mRNA. . .
N=952, Prob=1.6e-69; TR:P92188 P92188 PROTEASOME 29 KD SUBUNIT
(BC. . . N=551, Prob=1.2e-78, genomic survey sequence.
AL354405
AL354405.1 GI:7672051
GSS.

```

```

SOURCE      Leishmania major
            Leishmania major
            Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.

```

```

REFERENCE   1 (bases 1 to 537)
AUTHORS    Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and
            Smith,D.E.
TITLE      A physical map of the Leishmania major Friedlin genome
JOURNAL    Genome Res. 8 (2), 135-145 (1998)
MEDLINE    98146435
PubMed     9477341

```

```

REFERENCE   2 (bases 1 to 537)
AUTHORS    Taylor,R.G., Huckle,E.E.J., Ivens,A.C., Rajandream,M.A. and
            Barrell,B.G.
TITLE      Direct Submission

```

```

JOURNAL    Submitted (28-APR-2000) Leishmania major Friedlin genome sequencing
            Project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            alicat@sanger.ac.uk
            see http://www.ebi.ac.uk/parasites/leish.html
            details of Leishmania sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/L_major/
            The cDNA t7 primer sequence can be obtained from acc. no. US9231.

```

FEATURES

```

source
1..537
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="cosmid L8007.2"

```

ORIGIN

```

Query Match      5.0%; Score 21; DB 29; Length 537;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 156 GGACGAGAGAGTGTCTGCA 176

DB 397 GGACGAGAGAGTGTCTGCA 417

```

RESULT 14
A2829378      576 bp  DNA      linear  GSS 20-FEB-2001
LOCUS      2M010706F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCG2M010706 F, genomic survey sequence.

```

ACCESSION A2829378.1 GI:12999286

VERSION GSS.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

```

REFERENCE   1 (bases 1 to 576)
AUTHORS    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D., Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std. Error: 0.00
            Plate: 0107 row: F column: 06
            Seq primer: CGTGTAAACGACGCCAGT
            Class: plasmid ends
            High quality sequence stop: 576.

```

FEATURES

```

source
1..576
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M010706"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/notes="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptorized DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g14732114[gbl/AP129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptorized mouse DNA was annealed to
adaptorized vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

ORIGIN

```

Query Match      5.0%; Score 21; DB 28; Length 576;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 187 CTCACCAACAGCGCATGCGG 207

Db 232 CTCACCCACGCGCCATGCGG 212

RESULT 15

B1649379

LOCUS

DEFINITION

B1649379 655 bp mRNA linear EST 12-SEP-2001
603277822P1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5318165 5',
mRNA sequence.

ACCESSION

B1649379

VERSION

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 655)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLAM1806 row: 1 column: 06
High quality sequence start: 4
High quality sequence stop: 653.
Location/Qualifiers

FEATURES

source

/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5318165"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

ORIGIN

Query Match

5.0%; Score 21; DB 12; Length 655;

Best Local Similarity

100.0%; Pred. No. 87;

Matches 21; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY 115 CTCACCCCTACCTGCGCCAG 135

DB 442 CTCACCCCTACCTGCGCCAG 462

Search completed: March 22, 2004, 03:12:48
Job time: 3173 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 00:35:25 ; Search time 322 Seconds
(without alignments)
4790.519 Million cell updates/sec

Title: US-10-032-159A-15

Perfect score: 417
Sequence: 1 atcgggggaactgtgcgcag.....ggctcttgtaaccacgaa 417

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2438257 seqs, 1849576744 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	417	100.0	417	US-10-032-159A-15	Sequence 15, Appl
2	415	99.5	3417	US-09-767-215-4	Sequence 7, Appl
3	351	84.2	3012	US-09-767-215-3	Sequence 3, Appl
4	351	84.2	3766	US-10-104-047-315	Sequence 315, App
5	351	84.2	3931	US-09-767-215-1	Sequence 1, Appl
6	276	26.2	276	US-10-032-159A-17	Sequence 17, Appl
7	110	26.4	1141	US-10-032-159A-37	Sequence 37, Appl
8	22	5.3	509	US-10-027-632-280760	Sequence 280760,
9	22	5.3	509	US-10-027-632-280761	Sequence 280761,
10	20	4.8	534	US-10-282-122A-25923	Sequence 25923, A
11	19	4.6	867	US-09-821-831-3	Sequence 3, Appl
12	19	4.6	3096	US-09-798-917-9	Sequence 9, Appl
13	19	4.6	3096	US-10-325-917-9	Sequence 9, Appl
14	19	4.6	3259	US-10-388-934-557	Sequence 557, App
15	19	4.6	3260	US-09-821-831-1	Sequence 1, Appl

16	19	4.6	3949	US-09-798-412-7	Sequence 7, Appl
17	19	4.6	3949	US-10-325-917-7	Sequence 7, Appl
18	19	4.6	5790	US-10-134-102-7	Sequence 10, Appl
19	19	4.6	1163020	US-10-398-221-10	Sequence 10, Appl
20	19	4.6	3011208	US-10-398-221-2058	Sequence 2058, Ap
21	18	4.3	48	US-10-079-802-19	Sequence 19, Appl
22	18	4.3	48	US-10-016-986-35	Sequence 35, Appl
23	18	4.3	48	US-10-273-973-107	Sequence 107, App
24	18	4.3	708	US-10-016-986-40	Sequence 40, Appl
25	18	4.3	708	US-10-273-973-113	Sequence 113, App
26	18	4.3	908	US-10-424-599-82332	Sequence 82332, A
27	18	4.3	984	US-10-282-122A-31670	Sequence 31670, A
28	18	4.3	1830	US-10-156-761-7245	Sequence 7245, Ap
29	18	4.3	2390	US-10-221-278-81	Sequence 81, Appl
30	18	4.3	2390	US-10-119-926-35	Sequence 35, Appl
31	18	4.3	2390	US-10-291-172-81	Sequence 81, Appl
32	18	4.3	4054	US-10-321-195-9	Sequence 9, Appl
33	18	4.3	7654	US-10-006-773-1	Sequence 1, Appl
34	18	4.3	7654	US-10-006-771A-1	Sequence 1, Appl
35	18	4.3	41100	US-09-755-665-46	Sequence 46, Appl
36	18	4.3	9025608	US-10-156-761-1	Sequence 1, Appl
37	17	4.1	60	US-09-908-975-11168	Sequence 11168, A
38	17	4.1	130	US-09-994-228-53	Sequence 53, Appl
39	17	4.1	130	US-09-994-228-54	Sequence 54, Appl
40	17	4.1	252	US-10-156-761-5536	Sequence 5536, Ap
41	17	4.1	272	US-09-867-701-5725	Sequence 5725, Ap
42	17	4.1	295	US-10-424-599-117090	Sequence 117090,
43	17	4.1	362	US-09-864-761-33007	Sequence 33007, A
44	17	4.1	385	US-09-918-995-37596	Sequence 37596, A
45	17	4.1	435	US-10-425-114-20661	Sequence 20661, A

ALIGNMENTS

RESULT 1
US-10-032-159A-15
; Sequence 15, Application US/10032159A
; Publication No. US20020164703A1
; GENERAL INFORMATION:
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
; FILE REFERENCE: P-LJ 5100
; CURRENT APPLICATION NUMBER: US/10/032.159A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,457
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(417)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 416..417
; OTHER INFORMATION: n = A,T,C or G
US-10-032-159A-15

Query Match 100.0%; Score 417; DB 13; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.3e-209;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 ATCGGGGAAGTGTGCGGAGGACTCGGCACTGACGGAGTGAAGAGAGACTGTGG 60
1 ATCGGGGAAGTGTGCGGAGGACTCGGCACTGACGGAGTGAAGAGAGACTGTGG 60

Qy	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Qy	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Dy	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Qy	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
Dy	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
Qy	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220
Dy	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220
Qy	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280
Dy	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280
Qy	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340
Dy	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340
Qy	361	362	363	364	365	366	367	3																																

RESULT 2
US-09-767-215-4
; Sequence 4, Application US/09767215
; Patent No. US20020081636A1

```

1 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
2 TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
3
4 FILE REFERENCE: 0734-4-142001
5
6 CURRENT APPLICATION NUMBER: US/09/767,215
7
8 CURRENT FILING DATE: 2001-01-22
9
10 PRIOR APPLICATION NUMBER: 60/181,159
11
12 PRIOR FILING DATE: 2000-02-09
13
14 NUMBER OF SEQ ID NOS: 10
15
16 SOFTWARE: PaateSEO for Windows Version 4.0
17
18 SEQ ID NO 4
19
20 LENGTH: 3417
21
22 TYPE: DNA
23
24 ORGANISM: Homo sapiens
25
26 FEATURE:
27
28 NAME/KEY: CDS
29
30 LOCATION: (1)...(3417)
31
32 OS-09-767-215-4

```

Query Match	99.5%;	Score 415;	DB 9;	Length 3417;
Best Local Similarity	100.0%;	Pred. No. 1.1e-208;		
Matches 415;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGGGGAACTGTGCGCAGGGATCTCCGACTCAAGGACATGACAGAGACACTGTGG	60
QY	1	ATGGGGAACTGTGCGCAGGGATCTCCGACTCAAGGACATGACAGAGACACTGTGG	60
Db	1	ATGGGGAACTGTGCGCAGGGATCTCCGACTCAAGGACATGACAGAGACACTGTGG	60
QY	61	GAGATGATGAGAGCCACCGCCACAGATCTGACGCTCATCTGCCCCACCGCTTACC	120
Db	61	GAGATGATGAGAGCCACCGCCACAGATCTGACGCTCATCTGCCCCACCGCTTACC	120
QY	121	CCCTAAGTGGCCAGGCGCCAGGTCGTGACGAGCTGACAGAGAGAGAGTCGACAGC	180
Db	121	CCCTAAGTGGCCAGGCGCCAGGTCGTGACGAGCTGACAGAGAGAGAGTCGACAGC	180
QY	181	CCCCGCTCAACCAACAGCGCCATGCGGGCCGGGCACTTGCTGATTGCTGAAGCTCGA	240
Db	181	CCCCGCTCAACCAACAGCGCCATGCGGGCCGGGCACTTGCTGATTGCTGAAGCTCGA	240
QY	241	GGGAAGACGGGGCCATTCGCTTCTGAGAGACCTGAAGTTCCACAACCTGAGGTTCAC	300
Db	241	GGGAAGACGGGGCCATTCGCTTCTGAGAGACCTGAAGTTCCACAACCTGAGGTTCAC	300

Qy	301	ACCCGTGTCACCGGGCTGAGAGCTATGTTTAACTTCAGTAACTTTAAGCGTGAAGAGCTCC	360
Db	301	ACCCGTGTCACCGGGCTGAGAGCTATGTTTAACTTCAGTAACTTTAAGCGGTGAAGAGCTCC	360
Qy	361	GACTTTGAACGGTTTGACGAGCACTTCTGAGAACTCCAGCGCTCCGAGTAAACCCAG	415
Db	361	GACTTTGAACGGTTTGACGAGCACTTCTGAGAACTCCAGCGCTCCGAGTAAACCCAG	415

RESULT 3
US-09-767-215-3

```

Sequence 3, accession US/076/215
Patent No. US0020081636a1
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE
TITLE OF INVENTION: PROTEIN FAMILY AND T
FILE REFERENCE: 0734-142001
CURRENT APPLICATION NUMBER: US/09/767,215
CURRENT FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/181,159
PRIOR FILING DATE: 2000-02-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 3012
TYPE: DNA
ORGANISM: Homo sapiens
US-09-767-215-3

```

Query Match	84.2%;	Score 351;	DB 9;	Length 3012;
Best Local Similarity	100.0%;	Pred. No. 6.3e-175;		
Matches 351;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	ATGGGGGAAC	CTGTG	CCCGCA	GGGAT	CTCCGCA	CTCA	CGGAT	CTTGGAC	CGAGAGAC	CACTGTGG	60
Db	1	ATGGGGGAAC	CTGTG	CCCGCA	GGGAT	CTCCGCA	CTCA	CGGAT	CTTGGAC	CGAGAGAC	CACTGTGG	60
Qy	61	GAGATGATGAG	AGCCAC	CGGCCA	GAGAT	TCGAT	CGCTCAT	CTCCGCA	GGCGCT	CACC	120	
Db	61	GAGATGATGAG	AGCCAC	CGGCCA	GAGAT	TCGAT	CGCTCAT	CTCCGCA	GGCGCT	CACC	120	
Qy	121	CCCTACCTGCG	CAGGCCA	AGGTG	CTGTG	CCAGCTTGA	CGAGAGAG	AGTGTG	CTCACA	CAGC	180	
Db	121	CCCTACCTGCG	CAGGCCA	AGGTG	CTGTG	CCAGCTTGA	CGAGAGAG	AGTGTG	CTCACA	CAGC	180	
Qy	181	CCCCGGCTCA	CAACAG	CGCCAT	TGCGGGC	CGGGACA	CTTGCTGA	ATTTCGTGA	AAGCTTGA	240		
Db	181	CCCCGGCTCA	CAACAG	CGCCAT	TGCGGGC	CGGGACA	CTTGCTGA	ATTTCGTGA	AAGCTTGA	240		
Qy	241	GGGAAAGAA	GGGGCCAT	TGCGCTT	CCTGGA	AGAGCTTGA	AGTTTCCA	CAACCCCTGA	AGTCTTAC	300		
Db	241	GGGAAAGAA	GGGGCCAT	TGCGCTT	CCTGGA	AGAGCTTGA	AGTTTCCA	CAACCCCTGA	AGTCTTAC	300		
Qy	301	ACCGTGTGTA	CGGGTGC	AGCCAT	ATGTGA	CTTCA	CTTA	CTTTAG	GGGCT	351		
Db	301	ACCGTGTGTA	CGGGTGC	AGCCAT	ATGTGA	CTTCA	CTTA	CTTTAG	GGGCT	351		

RESULT 4
US-10-104-047-315

```

1 Sequence 315, Application US/10104047
2 Publication No. US20030236592A1
3
4 GENERAL INFORMATION:
5
6 APPLICANT: HELIX RESEARCH INSTITUTE
7
8 TITLE OF INVENTION: No. US20030236592A1e1 full length cDNA
9
10 FILE REFERENCE: H1-A0105
11
12 CURRENT APPLICATION NUMBER: US/10/104,047
13
14 CURRENT FILING DATE: 2002-03-25
15
16 PRIOR APPLICATION NUMBER:
17
18 PRIOR FILING DATE:
19
20 NUMBER OF SEQ ID NOS: 4096
21
22 SOFTWARE: PatentIn Ver. 2.1.1

```


SEQ ID NO 315
LENGTH: 3766
TYPE: DNA
ORGANISM: Homo sapiens
US-10-104-047-315

Query Match
Best Local Similarity 100.0%; Score 351; DB 15; Length 3766;
Pred. No. 6.2e-175;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGGAGGAACTCCGCACTCAGGCACTGAGAGAGACACTGTGG 60
DB 192 ATGGGGGAACTGTGCGGAGGAACTCCGCACTCAGGCACTGAGAGAGACACTGTGG 251
QY 61 GAGATGATGAGAGCCACCGCCACAGAGATGTAAGCTGATGTGCCACCGCCTCAC 120
DB 252 GAGATGATGAGAGCCACCGCCACAGAGATGTAAGCTGATGTGCCACCGCCTCAC 311
QY 121 CCTTACTGCGGCGGAGGCAAGGCTGTCGCAAGTGAAGAGAGAGAGTGTGCAAGC 180
DB 312 CCTTACTGCGGCGGAGGCAAGGCTGTCGCAAGTGAAGAGAGAGTGTGCAAGC 371
QY 181 CCCCGGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGCTGCA 240
DB 372 CCCCGGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGCTGCA 431
QY 241 GGGAGAACGGGGCCATGCGCTTCTGAGAGAGCTGAAGTTCCAGAACCTGACGCTTAC 300
DB 432 GGGAGAACGGGGCCATGCGCTTCTGAGAGAGCTGAAGTTCCAGAACCTGACGCTTAC 491
QY 301 ACCCTGTGTCACCGGGGCTGCAAGCTGATGTTGACTTCAAGTAACCTTAAGCGGT 351
DB 492 ACCCTGTGTCACCGGGGCTGCAAGCTGATGTTGACTTCAAGTAACCTTAAGCGGT 542

RESULT 5

US-09-767-215-1
Sequence 1, Application US/09767215
Patent No. US2002081636A1
GENERAL INFORMATION:

APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE REFERENCE: 07334-142001
CURRENT APPLICATION NUMBER: US/09/767,215
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/181,159
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 3931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (207)...(3218)
US-09-767-215-1

Query Match
Best Local Similarity 100.0%; Score 351; DB 9; Length 3931;
Pred. No. 6.1e-175;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGAACTGTGCGGAGGAACTCCGCACTCAGGCACTGAGAGAGACACTGTGG 60
DB 207 ATGGGGGAACTGTGCGGAGGAACTCCGCACTCAGGCACTGAGAGAGACACTGTGG 266
QY 61 GAGATGATGAGAGCCACCGCCACAGAGATGTAAGCTGATGTGCCACCGCCTCAC 120
DB 267 GAGATGATGAGAGCCACCGCCACAGAGATGTAAGCTGATGTGCCACCGCCTCAC 326
QY 121 CCTTACTGCGGCGGAGGCAAGGCTGTCGCAAGTGAAGAGAGAGTGTGCAAGC 180

DB 327 CCTTACTGCGGCGGAGGCAAGGCTGTCGCAAGTGAAGAGAGTGTGCAAGC 386
QY 181 CCCCGGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGCTGCA 240
DB 387 CCCCGGCTCAGCAACAGCGCCATGCGGCGCGGCACTTGCTGATTTGCTGAAGCTGCA 446
QY 241 GGGAGAACGGGGCCATGCGCTTCTGAGAGAGCTGAAGTTCCAGAACCTGACGCTTAC 300
DB 447 GGGAGAACGGGGCCATGCGCTTCTGAGAGAGCTGAAGTTCCAGAACCTGACGCTTAC 506
QY 301 ACCCTGTGTCACCGGGGCTGCAAGCTGATGTTGACTTCAAGTAACCTTAAGCGGT 351
DB 507 ACCCTGTGTCACCGGGGCTGCAAGCTGATGTTGACTTCAAGTAACCTTAAGCGGT 557

RESULT 6

US-10-032-159A-17
Sequence 17, Application US/10032159A
Publication No. US20020164703A1
GENERAL INFORMATION:

APPLICANT: Pawlowski, Krzysztof
APPLICANT: Reed, John C.
TITLE OF INVENTION: CARD-DOMAIN CONTAINING POLYPEPTIDES,
TITLE OF INVENTION: ENCODING NUCLEIC ACIDS, AND METHODS OF USE
FILE REFERENCE: P-LJ 5100
CURRENT APPLICATION NUMBER: US/10/032,159A
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,457
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 276
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(276)
US-10-032-159A-17

Query Match
Best Local Similarity 100.0%; Score 276; DB 13; Length 276;
Pred. No. 2.9e-135;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GAGGAGACACTGTGAGGAGATGATGAGAGAGCCACCGCCACAGATGTAAGCTGATGTC 105
DB 1 GAGGAGACACTGTGAGGAGATGATGAGAGAGCCACCGCCACAGATGTAAGCTGATGTC 60
QY 106 CCCAGCGGCTCACCCCTACTGCGGCGGAGGCAAGGCTGTCGCAAGTGAAGAGAG 165
DB 61 CCCAGCGGCTCACCCCTACTGCGGCGGAGGCAAGGCTGTCGCAAGTGAAGAGAG 120
QY 166 GAGGTGCTGCAAGCCCGGCTCACCAGAGGCGCATGCGGCGCGGCACTTGCTGAT 225
DB 121 GAGGTGCTGCAAGCCCGGCTCACCAGAGGCGCATGCGGCGCGGCACTTGCTGAT 180
QY 226 TTGCTGAAGACTGAGGAGAGAAAGGCGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCAC 285
DB 181 TTGCTGAAGACTGAGGAGAGAAAGGCGGCGCATGCGCTTCTGAGAGAGCTGAAGTTCCAC 240
QY 286 AACCTGACGTCACCTGTCGCAAGCTGTCACCGGGGCTGCAAG 321
DB 241 AACCTGACGTCACCTGTCGCAAGCTGTCACCGGGGCTGCAAG 276

RESULT 7

US-10-032-159A-37
Sequence 37, Application US/10032159A
Publication No. US20020164703A1
GENERAL INFORMATION:

APPLICANT: Pawlowski, Krzysztof
APPLICANT: Reed, John C.

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 76614
SOFTWARE: Patentin version 3.1
SEQ ID NO 25923
LENGTH: 534
TYPE: DNA
ORGANISM: Mycobacterium avium
US-10-282-122A-25923

Query Match 4.8%; Score 20; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 TGGTCACCGGCTGCAGCCT 324
|||||
Db 290 TGGTCACCGGCTGCAGCCT 309

RESULT 11
US-09-821-831-3/c
Sequence 3, Application US/09821831
Patent No. US20020137188A1
GENERAL INFORMATION:
APPLICANT: Bartlett, Perry Francis
APPLICANT: Coulson, Elizabeth Jane
APPLICANT: Fieldew, Katrina
APPLICANT: Baca, Manuel
APPLICANT: Kilpatrick, Trevor
APPLICANT: Surindar, Chema
TITLE OF INVENTION: Method of Modulating Cell Survival and
FILE REFERENCE: 3206,1001-000
CURRENT APPLICATION NUMBER: US/09/821,831
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: PCT/AU99/00860
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: AU P60701
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: AU P6351
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: AU P6353
PRIOR FILING DATE: 1998-10-06
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 867
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic CDNA sequence of rat
NAME/KEY: CDS
LOCATION: (115)...(867)
US-09-821-831-3

Query Match 4.8%; Score 19; DB 9; Length 867;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 AGGAACCTCAGGCTCCTGG 406
|||||
Db 771 AGGAACCTCAGGCTCCTGG 753

RESULT 12
US-09-798-412-9

Sequence 9, Application US/09798412
Publication No. US20030109428A1
GENERAL INFORMATION:
APPLICANT: Berlin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE REFERENCE: 07334-327001
CURRENT APPLICATION NUMBER: US/09/798,412
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 09/728,260
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: US 09/685,791
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/513,904
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/507,533
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: US 60/168,780
PRIOR FILING DATE: 1999-12-03
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 3096
TYPE: DNA
ORGANISM: Homo sapiens
US-09-798-412-9

Query Match 4.6%; Score 19; DB 10; Length 3096;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 GGACGAGGAGAGGTGCTG 174
|||||
Db 180 GGACGAGGAGAGGTGCTG 198

RESULT 13
US-10-325-917-9
Sequence 9, Application US/10325917
Publication No. US20030113787A1
GENERAL INFORMATION:
APPLICANT: Berlin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE REFERENCE: 07334-327001
CURRENT APPLICATION NUMBER: US/10/325,917
CURRENT FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US/09/798,412
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 09/728,260
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: US 09/685,791
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/513,904
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/507,533
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: US 60/168,780
PRIOR FILING DATE: 1999-12-03
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 3096
TYPE: DNA
ORGANISM: Homo sapiens
US-10-325-917-9

Query Match 4.6%; Score 19; DB 14; Length 3096;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 GGACGAGGAGAGGTGCTG 174
|||||

DB 180 GGACGAGAGAGGTGCTG 198

RESULT 14

US-10-388-934-557/c
; Sequence 557, Application US/10388934
; Publication No. US20040005547A1

GENERAL INFORMATION:

APPLICANT: Boess, Franziska
APPLICANT: Suter-Dick, Laura
APPLICANT: Wolf, Detlef

TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
FILE REFERENCE: 21169

CURRENT APPLICATION NUMBER: US/10/388,934
CURRENT FILING DATE: 2003-03-14

PRIOR APPLICATION NUMBER: 02005336.9
PRIOR FILING DATE: 2002-03-14

PRIOR APPLICATION NUMBER: 02015657.6
PRIOR FILING DATE: 2002-07-17

NUMBER OF SEQ ID NOS: 862
SOFTWARE: PatentIn version 3.1

SEQ ID NO 557
LENGTH: 3259

TYPE: DNA

ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-557

Query Match

4.6%; Score 19; DB 15; Length 3259;

Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 388 AGGAACTCAGGCTCCTGG 406

DB 770 AGGAACTCAGGCTCCTGG 752

RESULT 15

US-09-821-831-1/c

Sequence 1, Application US/09821831
Patent No. US20020137188A1

GENERAL INFORMATION:

APPLICANT: Bartlett, Perry Francis
APPLICANT: Coulson, Elizabeth Jane
APPLICANT: Fieldew, Katrina

APPLICANT: Baca, Manuel
APPLICANT: Kilpatrick, Trevor

APPLICANT: Surindar, Cheema
TITLE OF INVENTION: Method of Modulating Cell Survival and
TITLE OF INVENTION: Reagents useful for Same

FILE REFERENCE: 3206.1001-000
CURRENT APPLICATION NUMBER: US/09/821,831

CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: PCT/AU99/00860

PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: AU P00701

PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: AU PP6351

PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: AU PP6353

PRIOR FILING DATE: 1998-10-06
NUMBER OF SEQ ID NOS: 72

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1

LENGTH: 3260

TYPE: DNA

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Synthetic CDNA Sequence of Human
NAME/KEY: CDS

LOCATION: (115)...(1389)
US-09-821-831-1

Query Match

4.6%; Score 19; DB 9; Length 3260;

Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

388 AGGAACTCAGGCTCCTGG 406

DB

771 AGGAACTCAGGCTCCTGG 753

Search completed: March 22, 2004, 03:18:38
Job time : 344 secs